Copyright

by

Jennifer Ann Hennigan

2013

The Dissertation Committee for Jennifer Ann Hennigan Certifies that this is the approved version of the following dissertation:

Insights into the regulation of the DEAH-box helicase Prp43p through its interactions with three G-patch proteins

Committee:

Scott Stevens, Supervisor

Karen Browning

Jon Huibregtse

Arlen Johnson

Christopher Sullivan

Insights into the regulation of the DEAH-box helicase Prp43p through its interactions with three G-patch proteins

by

Jennifer Ann Hennigan, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2013

Dedication

For my husband Sean, who is the picture of Love illustrated in my life every day.

To my parents, who brought me up in the way I should go,

and I have not departed from it.

Acknowledgements

I would like to express my gratitude to my graduate supervisor Scott Stevens for allowing me to work with him for over six years. I appreciate his guidance in my research and kindness towards me throughout my tenure as a graduate student. I would also like to thank the members of my dissertation committee, Dr. Karen Browning, Dr. Jon Huibregtse, Dr. Arlen Johnson, and Dr. Chris Sullivan, for their helpful inights into my research. I express my gratitude for all members of the Stevens lab, past and present. In particular: Josh Combs, who was instrumental in my training as a graduate student; Rea Lardelli, Champ Gupton, and Amir Alpret, for engaging converstations regarding my research and life; Christine Cucinotta, who I had the privilege to mentor and was instrumental in the genetic screen for *prp43* mutants; Matt Sorenson, Sujin Lee, and Sanjeev Namjoshi, for creating a happy work place; and Al MacKrell, who has invested in my professional development and has provided scientific advice. I appreciate all of your frendships. Additionally, I am grateful for all the members of the Johnson laboratory who have always been generous with resources and scientific insights. I am also thankful for my friendship with Nancy Lyon, who has helped me with mammalian work and always brings me joy. I would be lost without all the love and support of my wonderful husband, my family, and my friends in Austin.

Insights into the regulation of the DEAH-box helicase Prp43p through its interactions with three G-patch proteins

Jennifer Ann Hennigan, Ph.D.

The University of Texas at Austin, 2013

Supervisor: Scott Stevens

The RNA helicase Prp43p is one of the few members of the DEAH-box helicase family that is known to operate in more than one cellular process in *Saccharomyces cerevisiae*. With roles in ribosome biogenesis and pre-mRNA splicing, Prp43p may be important in maintaining a communication conduit between these two pathways. Our studies provide insights into how Prp43p function is regulated through the use of three cofactors, Ntr1p, Pfa1p, and Gno1p, all of which interact with Prp43p at different steps of pre-mRNA splicing or ribosome biogenesis. Each cofactor contains a unique G-patch domain and our data show that they associate with Prp43p in a mutually exclusive manner. A strong growth defect and RNA processing phenotypes are seen upon overexpression of Pfa1p due to the dominance of Pfa1p interaction with Prp43p. Moreover, excess Pfa1p precludes Prp43p from interacting with either 35S pre-rRNA or U6 snRNA, indicating this one cofactor can negatively regulate Prp43p recruitment into ribosome biogenesis and pre-mRNA splicing pathways, respectively. We have determined that Ntr1p and Gno1p are able to compete with one anther for Prp43p

occupancy. Similar to Ntr1p, we show that the G-patch domain of Gno1p contributes to its association with Prp43p.

To further understand pathway specificity of Prp43p, we characterized conditional *prp43* alleles with mutations C-terminal to the conserved RecA domains of Prp43p. These novel alleles affect pre-mRNA splicing and ribosome biogenesis, though none are mutually exclusive. Multiple *prp43* alleles are deficient in tri-snRNP formation, a previously uncharacterized phenotype in *prp43* mutants. The majority of our *prp43* mutants display varying rRNA defects, with some alleles impacting ribosome biogenesis more severely or moderately than known *prp43* ATPase mutants. To correlate the processing defects seen in each allele, we have determined the extent of association of the mutants with each G-patch protein. Altogether, our data support a working model for Prp43p in which its substrate specificity, activation, and cellular distribution is coordinated through the efforts of the three G-patch proteins in yeast and sheds light on potential mechanisms of general DExH/D helicase function and regulation.

Table of Contents

List of Tables

List of Figures

List of Illustrations

Chapter 1: Introduction

1.1 Eukaryotic gene expression

In eukaryotes, efficient gene expression is dependent upon RNA and proteins complexed together in several ribonucleoprotein (RNP) machineries. From the moment transcription by RNA Polymerase II (RNA Pol II) begins in the nucleus, a pre-mRNA is bound by heterologous nuclear RNP (hnRNP) proteins and serine/arginine-rich (SR) proteins (rev. in Wahl *et al*. 2009). While in the nucleus, intervening sequences in pre-mRNAs called introns are removed through the action of the spliceosome, a multiple mega-dalton complex composed of five small nuclear RNPs (snRNPs) and dozens of trans-acting factors. Each snRNP is comprised of a small nuclear RNA (snRNA), either U1, U2, U4, U5 or U6, and both common and specific proteins (rev in Hoskins and Moore 2012). Mature mRNAs are exported through the nuclear pore complex (NPC) to the cytoplasm and are translated by another large RNP machine, the ribosome (rev. in Mor and Shav-Tal 2010). The ribosomal RNAs (rRNA) that compose the ribosome often have modifications that are directed by small nucleolar RNAs (snoRNA) complexed with proteins as small nucelolar RNPs (snoRNP) (rev. in Watkins and Bohnsack 2012). Cells also use other non-coding RNAs (ncRNA) like microRNAs (miRNAs) to provide more subtle regulation over gene expression. miRNAs can affect the stability or translation of a target mRNA through perfect or imperfect basepairing with the target, respectively. Through differential expression, miRNAs can regulate gene expression in a tissue- or developmental-stage-specific manner. Overall, all of these ncRNAs are essential in the infrastructure and the regulation of a eukaryotic cell (rev. in Mattick and Makunin 2006).

1.2 PRE-MRNA SPLICING

The spliceosome works to remove non-coding sequences, termed introns, from pre-mRNAs. In the process, the coding sequences, termed exons, are ligated to form a mature substrate for translation. The snRNP complexes that make up the major spliceosome contain the U1, U2, U4, U5 and U6 snRNAs, as well as both common and unique proteins. Most pre-mRNAs are addressed by the major spliceosome, while in higher eukaryotes, a smaller number of transcripts are processed via the minor spliceosome. This complex also contains the U5 snRNP, but the U1, U2, U4, and U6 snRNPs are substituted by the U11, U12, U4atac, and U6atac snRNPs, respectively. Since the minor spliceosome is only present in higher eukaryotes and is thought to generally act in a similar manner to the major, it will not be extensively covered here (rev. in Patel and Bellini 2008). My work presented here was conducted in *Saccharomyces cerevisiae*, but since 85% of yeast pre-mRNA splicing factors have clear orthologs in humans, the yeast spliceosome is believed to represent the core of the human spliceosome (Fabrizio *et al*. 2009).

1.2.1 Functional importance of introns

The prevalence and length of introns appears to vary between species. In yeast, only 5% of genes contain introns within some 283 genes (Spingola *et al*. 1999; Juneau *et al*. 2007; Parenteau *et al*. 2008). Only a handful of these genes contain more than one intron, and most introns are relatively short (~100-400 bp) compared to human introns, which number on average seven per transcript and range to greater than 490,000 bp (Spingola *et al*. 1999; Sakharkar *et al*. 2004). Additionally, since human genes generally

contain multiple introns, different transcripts are commonly produced through alternative pre-mRNA splicing to include or exclude exons in the mRNA in a regulated manner (Nagasaki *et al*. 2005).

The introns of human genes also contain non-coding elements like snoRNAs and miRNAs (Rearick *et al*. 2011). For example, the mammalian U22 snoRNA gene contains ten introns, and seven other snoRNAs (U25-U31) encoded in its introns (Tycowski *et al*. 1996). In humans, snoRNAs are predominantly processed in a splicing-dependent manner. Generally, it is only after intron removal by the spliceosome and debranching of the intron lariat that the mature snoRNA is released. Human miRNAs can either be processed in a similar manner from the excised intron lariat, from the unspliced pre-mRNA, or concurrently with splicing. Such different pathways are hypothesized to provide competition between pre-mRNA splicing and miRNA biogenesis, which might regulate gene expression (Brown *et al*. 2008). In yeast, only six snoRNAs are known to be located in introns: U24, snR38, U18, snR44, snR59 and snR39 in *ASC1, TEF4, EBF1, RPS22B, RPL7B, RPL7A*, respectively. As in humans, these snoRNAs are dependent on Dbr1p to debranch the intron lariat for their release and proper maturation (Ooi *et al*. 1998). However, the loss of any of these snoRNAs does not affect cell growth in rich media (Parenteau *et al*. 2008, 2011). Only three genes in yeast contain introns that are important for cell growth under normal conditions (*MTR2*, *YRA1*, and *TAD3*), but the essentiality of these introns is dependent upon the promoter of the gene, indicating the introns are utilized for transcriptional control of the genes. However, the loss of some

introns can affect growth under stress such as hypo- or hyperosmotic conditions (Parenteau *et al*. 2008)

In yeast, 50% of RNA transcripts produced by RNA Pol II encode structural ribosomal proteins (r-proteins) (Ares *et al*. 1999; Warner 1999), and these transcripts comprise 90% of the intron-containing pre-mRNAs, making them the largest class of proteins transcribed by RNA Pol II and addressed by the spliceosome (Ares *et al*. 1999). A link between rRNA and mRNA processing has been hypothesized which may serve to maximize the regulation of ribosome production (Ares *et al*. 1999). Indeed, introns in r-protein genes appear to be involved in the regulation of ribosome biogenesis. If the introns of certain r-protein genes are removed, the expressions of the genes are affected. Additionally, within duplicated r-protein pairs, the majority of introns exert some intergenic regulation. For example, the removal of an intron in one r-protein paralog can reduce the expression of the other paralog. Therefore, some r-protein paralog pairs are expressed in an intron-dependent manner (Parenteau *et al*. 2011). The presence of introns in 15% of r-protein genes contributes to proper rRNA processing. Removal of these introns results in increased or decreased abundance in the 35S pre-rRNA levels. Increased 35S pre-rRNA levels, signifying a delay in processing, are specifically linked to introns of large subunit r-protein genes, several of which also affect A1 and D site cleavages (see Chapter 1.3.2 for details on cleavages) (Parenteau *et al*. 2011).

As additional evidence for coordination, the pre-mRNA and mRNA levels of r-proteins change in response to various environmental stresses (Pleiss *et al*. 2007; Bergkessel *et al*. 2011). For example, yeast respond to amino acid starvation by

decreasing the splicing efficiency of r-protein encoding pre-mRNAs in a transcript-specific manner, which results in pre-mRNA accumulation (Pleiss *et al*. 2007; Bergkessel *et al*. 2011). In contrast, stresses that lead to TORC1 inactivation, like heat shock and hyperosmotic stress, result in a rapid loss of r-protein pre-mRNAs rather than their accumulation. Therefore, pre-mRNA processing is a means by which the cell can modulate ribosome production appropriately in response to environmental conditions (Parenteau *et al*. 2011).

1.2.2 Pre-mRNA splicing mechanism

Pre-mRNAs are spliced within a 40S spliceosome in a series of ATP-dependent steps (Brody and Abelson 1985; Lin *et al*. 1987; Schwer and Gross 1998; Martin *et al*. 2002). First, the spliceosome must assemble on the pre-mRNA substrate, then it directs two transesterification reactions to join the exons. Finally, in a third step, the pre-mRNA and lariat intron are released (Lin *et al*. 1987). Two S_N2 transesterification reactions occur during splicing of a pre-mRNA (Illustration 1.1) (Domdey *et al*. 1984; Grabowski *et al*. 1984; Padgett *et al*. 1984; Rodriguez *et al*. 1984; Lin *et al*. 1987). First, the phosphodiester bond of the 5' splice site (5' SS) is nucleophilically attacked by the 2'-OH of a conserved adenosine in the branch point sequence within the intron (UACUA**A**C). This reaction yields the free 5' exon and the 3' exon attached to a circularized intron, called the lariat intermediate. The branchpoint adenosine is linked by a 2'-5'-phosphodiester bond to the guanosine from the conserved GU in the 5' SS (Breathnach *et al*. 1978), as well as the downstream cytosine by a typical 3'-5' phosphodiester bond (Domdey *et al*. 1984). In the second reaction, the 3'-OH of the 5' exon attacks the

phosphate following the conserved AG in the 3' splice site (3' SS) (Breathnach *et al*. 1978), joining the 3' and 5' exons. Unlike the self-splicing group II introns, pre-mRNAs cannot catalyze these two reactions alone, but depend on the *trans*-acting factors that comprise the spliceosome to fold and rearrange the pre-mRNA for catalysis. Additionally, none of the snRNAs of the spliceosome contain a pre-formed catalytic site. In fact, the RNA-RNA interactions within the spliceosome are generally weak and are dependent upon the protein factors for stability, scaffolding, and remodeling during the catalytic cycle (rev. in Wahl *et al*. 2009).

1.2.3 snRNP biogenesis

The U1-U5 snRNAs are transcribed by RNA Pol II and are capped co-transcriptionally with a 5'-7-methylguanosine cap (m^7G) like other RNA Pol II transcripts (rev. in Patel and Bellini 2008). In *S. cerevisiae*, snRNP biogenesis is completed within the nucleus, but in higher eukaryotes pre-snRNAs undergo a cytoplasmic maturation phase (Verheggen and Bertrand 2012). In both cases, a heteroheptameric ring of Sm proteins bind around a conserved sequence motif called the Sm site ($PuAU_{4-6}GPU$) (Raker *et al*. 1999). The methyltransferase Tgs1p recognizes the assembled Sm core and adds two methyl groups to the 5' cap to create a distinct hyper-methylated 2,2,7-trimethylguanosine cap (Verheggen and Bertrand 2012). In yeast, Tgs1p localizes to the nucleoli, indicating that pre-snRNPs transit through the nucleoli during maturation (Mouaikel *et al*. 2002). In humans, the pre-snRNAs are transported into the cytoplasm prior to Sm assembly in a CRM1-dependent manner through association with the Phosphorylated Adaptor for RNA export (PHAX). The formation of the Sm ring is

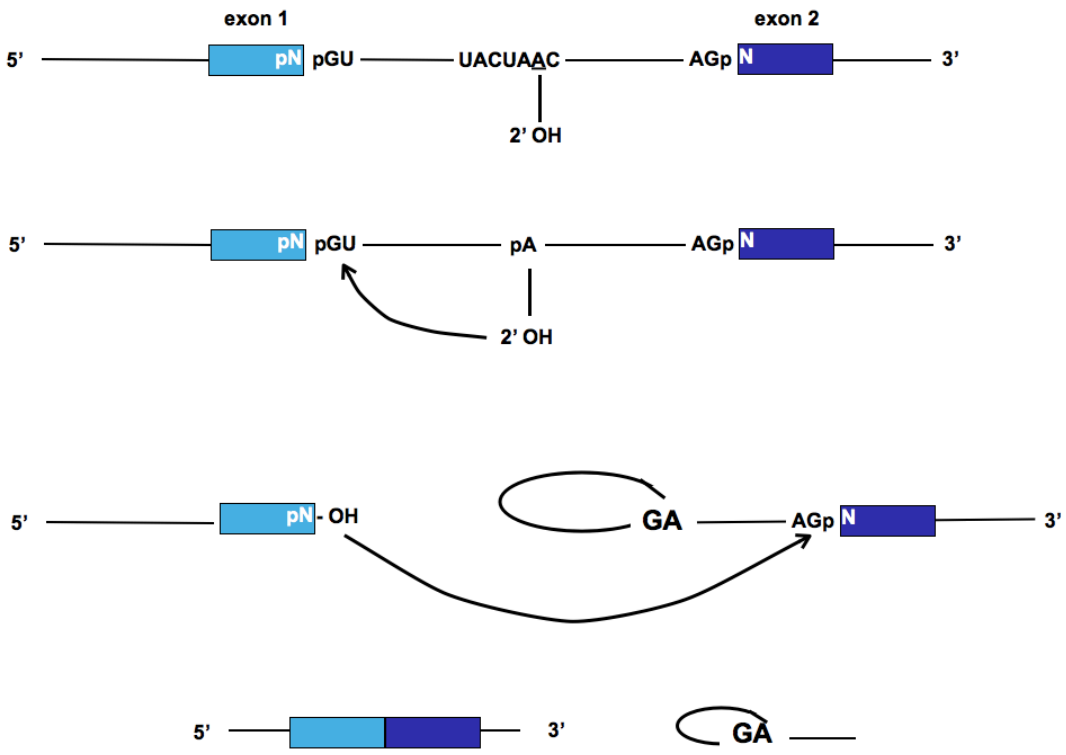


Illustration 1.1 Chemical reactions of pre-mRNA splicing.

Pre-mRNA splicing occurs through two transesterification reactions. First, the 2'OH of the branchpoint adenosine attacks the 5' SS to produce a free 5' exon and a lariat intermediate. The reactive adenosine in the concensus sequence is underlined. In the second reaction, the 5' SS attacks the conserved guanosine in 3'SS. The two exons are now ligated to create a mature mRNA and the intron is released as a branched 2'-5' phosphodiester lariat that is degraded.

facilitated by a multisubunit enzyme called the survival of motor neuron protein (SMN) complex. hTgs1p is a component of the SMN complex in the cytoplasm, and, following its activity, the trimethylated cap and the Sm core act as nuclear localization signals (NLS) (rev. in Patel and Bellini 2008). In metazoans, the adaptor protein Snurportin recognizes the trimethylated caps and enchances the trafficking of the nascent snRNPs into the nucleus for final maturation (Huber *et al*. 1998). In the nucleus, the snRNAs are modified with 2'-O-methylations and psueduridylations that are directed by small Cajal body-specific RNAs (scaRNAs) (Jády *et al*. 2003) and assembly with the remaining snRNP-specific proteins is completed therein.

In contrast to the other snRNAs, the U6 snRNA is transcribed by RNA Pol III. After transcription, the La protein binds both its 5'-triphosphate cap and a 3' U-rich motif, and protects the pre-snRNA from degradation (rev. in Patel and Bellini 2008). The 5' end of the U6 pre-snRNA is uniquely capped by a specific methyl transferase to produce a ϒ-monomethyl cap (Shimba and Reddy 1994). Additionally, the 3' end of the pre-snRNA is trimmed by a specific exonuclease and an unidentified enzyme produces a 2', 3' cyclic phosphate tail. Together, these two unique modifications preclude the continued binding of the La protein (rev. in Patel and Bellini 2008). Unlike other snRNPs, the U6 snRNP lacks the Sm site and proteins, but contains Sm-like proteins (Lsm 2-8). After removal of the La protein, the Lsm proteins bind to the same 3' U-rich motif as a heteroheptameric complex (Achsel *et al*. 1999). The Lsm proteins are necessary for retention of the U6 in the nucleus; therefore, the U6 pre-snRNP does not have a cytoplasmic maturation phase like other snRNPs in metazoa (Spiller *et al*. 2007). At some point during its maturation,

the U6 snRNA cycles through the nucleolus and snoRNAs direct 2'-O-ribose methylations and pseudouridylations (Ganot *et al*. 1999). After transitioning through the nucleolus, the U6 snRNA accumulates in Cajal bodies like the other snRNAs (Lange and Gerbi 2000). The mature U6 snRNP also contains the unique component Prp24p (SART3 in human) in addition to the Lsm core (Shannon and Guthrie 1991).

1.2.3.1 Di-snRNP biogenesis

The di-snRNP particle contains the U4 and U6 snRNAs extensively paired in two intermolecular stems (Brow and Guthrie 1988). This annealing is facilitated by the U6 RNA binding factor, Prp24p (Raghunathan and Guthrie 1998a). An intramolecular stem loop (ISL) in the 3' end of U6 must be disrupted for proper annealing of the U4 snRNA (Fortner *et al*. 1994). Four RNA recognition motifs (RRMs) in Prp24p bind the U6 snRNA surrounding the ISL (Martin-Tumasz *et al*. 2010, 2011). RRM1 and RRM2 bind along the stem of the ISL (Martin-Tumasz *et al*. 2010), while RRM3 binds closer to the loop (Martin-Tumasz *et al*. 2011). An occluded RRM, oRRM4, at the base of the helix is thought to be able to disrupt base pairing in the ISL. Prp24p is then thought to be released due to the more highly favorable free energy of the newly formed U4/U6 duplex. The intramolecular U4/U6 stem eliminates the binding site of Prp24p within the U6 snRNA (Martin-Tumasz *et al*. 2011). Due to the energetically favorable binding of U4/U6, little free U4 can be detected in wildtype conditions (Anthony *et al*. 1997; Martin-Tumasz *et al*. 2011). Consequently, the levels of free U6 snRNP is approximately five times greater than the levels of free U4 snRNP (Li and Brow 1993)

Due to the low abundance of free U4 snRNP, it has been difficult to determine the order of proteins that join the U4/U6 di-snRNP. However, in yeast, these di-snRNP proteins include Prp3p (Anthony *et al*. 1997), Prp4p (Banroques and Abelson 1989), and Prp31p (Weidenhammer *et al*. 1997). All of these have also been characterized as tri-snRNP components (Stevens and Abelson 1999) and are present in spliceosomes (Fabrizio *et al*. 2009). However, di-snRNP components are not found in activated spliceosomes, which indicates that they may leave coincident with U4 snRNA release (Fabrizio *et al*. 2009).

1.2.3.2 Tri-snRNP formation

The U4/U6 di-snRNP joins the 16S U5 snRNP containing Brr2p to form a 25S tri-snRNP species, U4/U6•U5. A complement of 24 proteins sediment with U4/U6•U5 (Stevens and Abelson 1999). Interactions between the U4/U6 and U5 snRNPs are mediated by proteins and not by RNA base pairing. Tri-snRNP is structured so that the U4/U6 snRNP extends from a centralized middle as an arm (Häcker *et al*. 2008). In yeast, Prp6p is essential for the formation of tri-snRNP (Galisson and Legrain 1993) and has been mapped in the linker region between the snRNPs. This linker also contains Prp31p, which lies adjacent to Prp6p. The heterodimer Prp3p/Prp4p is positioned further up the arm (Häcker *et al*. 2008). Prp3p also is necessary for tri-snRNP formation; its loss results in a reduced amount of tri-snRNP, most likely due to decreased stability of the U4/U6 di-snRNP *in vivo* (Anthony *et al*. 1997).

1.2.3.3 Penta-snRNP model

Data exists that supports two methods of spliceosome assembly, either a stepwise model or a pre-formed model. Under the low salt conditions required for *in vitro* pre-mRNA splicing, a 45S penta-snRNP can be isolated which contains all five snRNPs. This model postulates that the penta-snRNP binds the pre-mRNA substrates as a pre-formed complex, rather than individual steps of U1, U2, then tri-snRNP joining. Conspicuously, all the DEAH and DEAD-box helicases involved in pre-mRNA splicing are absent from the penta-snRNP (Stevens *et al*. 2002). Moreover, corresponding data has been found in HeLa and chicken cells with the discovery of the supraspliceosome (Chen *et al*. 2007). The supraspliceosome is a macromolecular complex consisting of four intact spliceosomes containing all five snRNPs. These four complexes are threaded with the pre-mRNA so that splice sites can be aligned and unprocessed regions of the pre-mRNA can be protected (Azubel *et al*. 2006). For the purpose of relating the regulation of pre-mRNA splicing, the details of the step-wise model will be outlined further and referenced throughout.

1.2.4 Spliceosome assembly and dynamics

In the step-wise model for yeast spliceosome assembly (Illustration 1.2) (rev. in Wahl *et al*. 2009; Hoskins and Moore 2012), the U1 snRNP recognizes the 5' SS and basepairs with it in an ATP-independent manner (Legrain *et al*. 1988; Séraphin *et al*. 1988). Additionally, the branchpoint binding protein (Msl5p/hBBP) and Mud2p bind to the branchpoint and a polypyrimidine tract that lies directly downstream of it. The pre-mRNA bound by the U1 snRNP and these proteins create the early complex (E complex)

or commitment complex (Hoskins and Moore 2012; Seraphin and Rosbash 1989). Past investigations indicated that the association of the U1 snRNP commits the pre-mRNA to complete splicing (Legrain *et al*. 1988), but more recent investigations have shown that pre-mRNAs can be released after U1 snRNP binding. Instead of a single event committing a pre-mRNA to splicing irreversibly, as the pre-mRNA passes quality control checks along the pathway, commitment to the splicing pathway increases (Hoskins *et al*. 2011; rev. in Cordin *et al*. 2012). Therefore, the term 'commitment complex' is misleading.

In an ATP-dependent reaction, Mud2p and Msl5p are replaced by U2 snRNP binding and basepairing to the branchpoint (rev. in Hoskins and Moore, 2012). The DEAD-box helicase Sub2p (hUAP56) is necessary for the removal of Mud2p, and possibly Msl5p; interestingly, the function of Sub2p can be bypassed by the deletion of Mud2p (Kistler and Guthrie 2001). Concurrently, Prp5p acts on the U2 snRNP in an ATP-dependent manner to remodel it so that its branchpoint binding region is accessible (Wiest *et al*. 1996). The addition of the U2 snRNP leads to the formation of the A complex. The SF3a and SF3b subcomplexes that are a part of the U2 snRNP help to stabilize the basepairing between the U2 snRNA and the intron (Gozani *et al*. 1996).

Next, the U4/U6•U5 tri-snRNP joins to make the B complex (or the pre-spliceosome in yeast), which is still catalytically inactive (rev. in Hoskins and Moore, 2012). Additionally, a large heteromeric complex of non-snRNP associated proteins called the Nineteen complex (NTC) associates at this time (Tarn *et al*. 1993). The B complex is activated through a series of RNA and protein rearrangements to create a

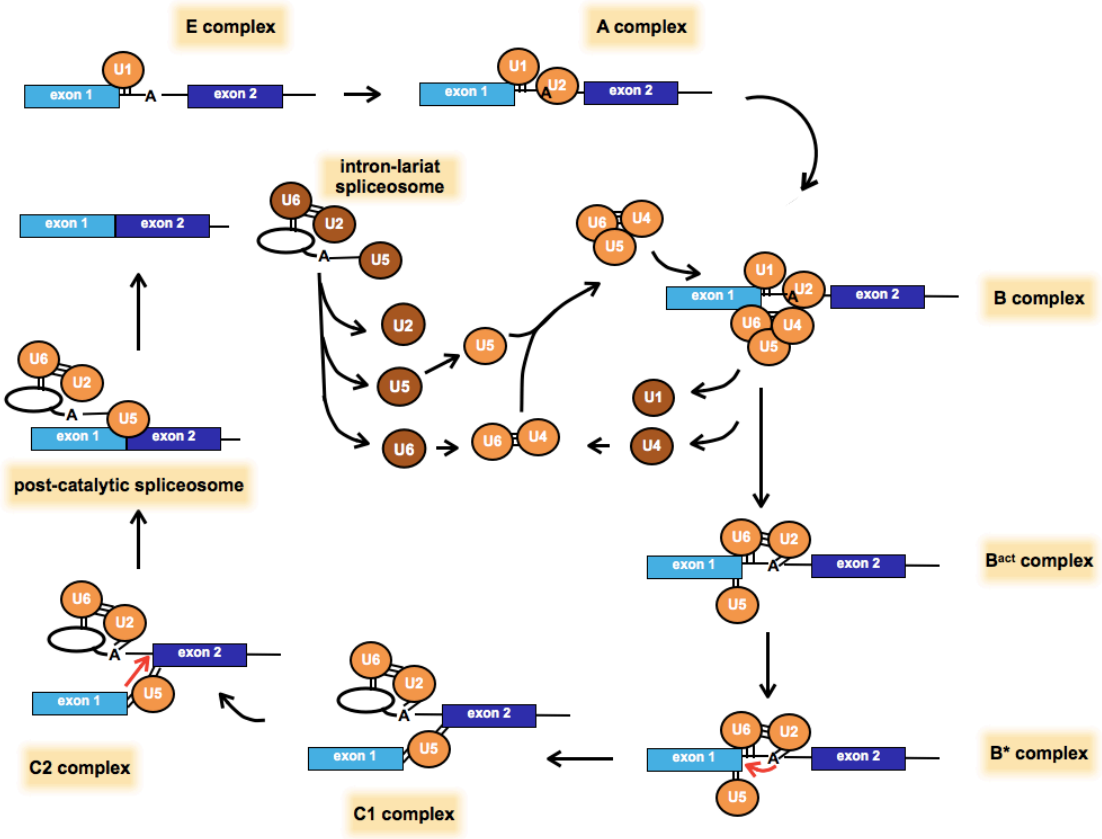


Illustration 1.2 Step-wise splicing cycle.

Pre-mRNAs can be addressed by the spliceosome in a step-wise manner (see Chapter 1.2.4 for details). The U1 snRNP first joins the pre-mRNA in a commitment complex, then the U2 snRNP basepairs with the brachpoint. Tri-snRNP joins and the spliceosome is activated through the removal of the U1 and U4 snRNPs. The two transesterification reactions are depicted by the red arrows. After splicing, the mature mRNA and lariat intron are released. SnRNPs that need to be recycled are depicted in dark orange. Di-snRNP is recycled from released U4 and U6 snRNPs (see Chapter 1.2.3.1) and U5 snRNP undergoes nuclear export and reentry for recycling (see Chapter 1.2.5.1).

catalytically active B^{act} complex (rev. in Hoskins and Moore, 2012). The U1 snRNP is released through the ATP-dependent action of Prp28p, which unwinds the U1:5' SS duplex (Staley and Guthrie 1999). Brr2p also acts in an ATP-dependent manner to unwind U4 from U6 (Raghunathan and Guthrie 1998b), which then allows U6 to substitute for U1 and allow baseparing between a conserved sequence (ACAGAG) and and the 5' SS (Lesser and Guthrie 1993). The U5 snRNP interacts with the 5' exon, and the NTC helps stabilize the interactions within the B^{act} complex. Without the NTC, the U6:5' SS interaction is not correctly positioned for spliceosome activation (Chan and Cheng 2005).

The Prp2p RNA helicase acts in an ATP-independent manner when it joins the spliceosome to induce a conformational change of the B^{act} complex. Then through an ATP-dependent action, Prp2p removes the SF3 subcomplexes, which has been proposed to expose the brachpoint adenosine (Lardelli *et al*. 2010). Now in the correct structural context, the B* spliceosome is formed and the first transesterification reaction proceeds (Lardelli *et al*. 2010). The C1 complex contains the free 5' exon and the lariat intermediate. To promote the next transesterification, the helicase Prp16p remodels the C1 complex into the C2 complex in an ATP-dependent manner (Schwer and Guthrie 1991). In the transition, the U6:5' SS basepairing must be destabilized to allow for the conformational change in the spliceosome that promotes the second step reaction. The 3'SS of the lariat-exon intermediate competes with the branched lariat intron for the active site. This competition between the first step products and the second step substrates can drive the reaction either forward or backward (Konarska *et al*. 2006; Tseng

and Cheng 2008). Two second step factors, Slu7p and Prp18p, help position the 3' SS correctly in the active site (James *et al*. 2002). Also, the DEAH-box helicase Prp22p joins the spliceosome before the second reaction and plays an ATP-independent role (Schwer and Gross 1998). Once in position in the C2 complex, the second transesterification reaction proceeds and the pre-mRNA splicing reaction is complete.

1.2.5 Spliceosome disassembly and recycling

After the splicing reaction, the post-catalytic spliceosome contains the U2/U6•U5 snRNPs still associated with the spliced mRNA and the lariat intron. Prp22p unwinds the duplex formed by loop 1 of the U5 snRNA and mRNA in an ATP-dependent manner to release the mature mRNA from the spliceosome (Schwer and Gross 1998; Schwer 2008). At the point of Prp22p action, the second step factors Slu7p and Prp18p are also removed (Schwer 2008). Slu7p and Prp18p have been proposed to be necessary for Prp22p to join the spliceosome (James *et al*. 2002); however, recent *in vitro* data indicate that these factors are not strictly required for Prp22p recruitment or function (Fourmann *et al*. 2013). Therefore, the removal of Slu7p and Prp18p is not the action that destabilizes the mRNA, but rather the helicase activity of Prp22p. In addition to the removal of Slu7p and Prp18p, Prp22p displaces the RES (Retention in splicing) complex as well as Cwc22p and Cwc21p (Fourmann *et al*. 2013). The RES complex functions to retain pre-mRNAs in the nucleus until splicing has been completed (Dziembowski *et al*. 2004), and its removal by Prp22p is thought to allow mRNAs to be exported from the nucleus (Fourmann *et al*. 2013). The significance of the removal of Cwc22 or Cwc21 is not yet known, but they are speculated to lie at the interface between U5 and the mRNA. The

intron-lariat spliceosome that remains after Prp22p action contains the U2, U5, and U6 snRNPs still complexed with with the lariat intron (Fourmann *et al*. 2013).

The intron lariat is removed from the snRNPs through the ATP-dependent action of another RNA helicase, Prp43p, which is recruited to the spliceosome as a component of the trimeric NTR complex with Ntr1p and Ntr2p (Martin *et al*. 2002). Previous data suggested that Brr2p works to unwind the U2/U6 basepairing similar to its role in activating the spliceosome (Small *et al*. 2006). However, recent *in vitro* data demonstrates that the Prp43p alone can completely separate the U2, U5, and U6 snRNPs without the assistance of Brr2p. A 20-25S U2 snRNP and an ~18S U5 snRNP were purified from *in vitro* splicing assays conducted with reconstituted spliceosomes after the action of Prp43p. U6 is released as a protein-free snRNA. There are no significant changes in the protein composition of the U2 or U5 snRNPs purified after Prp43p action compared to their compositions in the post-catalytic spliceosome. These data indicate that the U2 snRNP is not significantly remodeled during the splicing cycle (Fourmann *et al*. 2013), which stands in opposition to previous reports that U2 proteins are lost during catalysis (Lardelli *et al*. 2010) or after disassociation (Tsai *et al*. 2005), although it should be noted that those experiments were performed *in vivo* (Lardelli *et al*. 2010) vs. *in vitro* (Fourmann *et al*. 2013).

Following their removal from the spliceosome or disassembly, snRNPs are recycled for further rounds of splicing. As mentioned above, Prp24p plays an essential role in the recycling of the U4 and U6 snRNPs into di-snRNP (Raghunathan and Guthrie 1998a). The necessity to recycle the U2/U6•U5 post-catalytic snRNP *in vivo* is suggested

by the genetic interactions between the disassembly factor Prp43p and the tri-snRNP assembly factor Prp38p. Additionally, mutations in Ntr1p and the U5 recycling factor Aar2p can also suppress *prp38-1* mutants (Pandit *et al*. 2006). Though the U2 snRNP appears to be released intact, it is speculated that the U2 snRNA requires conformational changes induced by the DEAD-box RNA helicases Sub2p and Prp5p for its recycling. Additionally, the U5 snRNP that is released was only characterized as containing Brr2p, Snu114p, and Prp8 and must be recycled properly via Aar2p (see Chapter 1.2.51) (Fourmann *et al*. 2013).

Once the lariat intron in released, the debranching enzyme Dbr1p hydrolyzes the 2'-5' phosphodiester bond (Chapman and Boeke 1991). While Dbr1p is non-essential in *S. cerevisiae* and its deletion confers no growth defect, the loss of its homolog in *S. pombe* renders the cells slow growing and morphologically deformed (Chapman and Boeke 1991; Kim *et al*. 2000). These defects are thought to occur in *S. pombe* rather than *S. cerevisiae* due to the increased number of introns in the *S. pombe* genome. Likewise, it is hypothesized that the role of Dbr1p in higher eukaryotes is essential to clear lariats and release any ncRNAs that may be encoded within (Kim *et al*. 2000).

Additionally, stalled, aberrantly-formed splicing intermediates can also be disassembled through the action of Prp43p as a part of the NTR complex (Koodathingal *et al*. 2010; Chen *et al*. 2013). The interaction between the NTR complex and the spliceosome is mediated via Ntr2p and the C-terminal region of Brr2p comprising its second helicase domain and the C-terminus (aa 1209-2163). Since Brr2p is a stable component of the U5 snRNP, Ntr2p binding must be inhibited throughout the splicing

cycle to prevent premature disassembly. While Ntr2p can join a stalled splicesome at specific stages, its greatest affinity is for its natural substrate, the post-splicing lariat RNP (U2/U6•U5 plus lariat intron). The NTR complex can catalyze disassembly after the action of Prp2p or Prp16p, but not before or while these helicases are bound to the spliceosome. In fact, Prp16p can outcompete Ntr2p for binding to spliceosomes stalled at the Prp16-dependent step. Additionally, spliceosomal factors Cwc25p and Slu7p, involved at the first and second step, respectively, also prevent spliceosomes from being susceptible to disassembly. Altogether, the model suggests that aberrant pre-mRNAs, which are identified via kinetic proofreading, are substrates for disassembly after the ATP-dependent function of Prp2p, Prp16p and Prp22p (Chen *et al*. 2013).

1.2.5.1 U5 Recycling

In yeast, the U5 snRNP is released from the post-splicing lariat RNP as an 18S complex (Fourmann *et al*. 2013) which must be recycled into the 16S species which is incorporated into tri-snRNP (Gottschalk *et al*. 2001; Stevens *et al*. 2001; Tsai *et al*. 2007). In humans, the post-splicing U5 snRNP has been purified as a 35S particle and contains the NTC, which joins the U5 snRNP at spliceosome activation (Makarov *et al*. 2002). In yeast, only a small portion of released U5 snRNP contains the NTC, and a corresponding 35S species could not be recovered from *in vitro* resconstitution and disassembly assays (Fourmann *et al*. 2013) While the details are not completely understood regarding U5 snRNP recycling in yeast, the process appears to be essential *in vivo* (Gottschalk *et al*. 2001). One crucial step in this process involves the use and removal of recycling factor Aar2p (Weber *et al*. 2011).

The 16S U5 snRNP exists as two species, either with or without Aar2p (Gottschalk *et al*. 2001; Stevens *et al*. 2001). The Aar2p-U5 species is cytoplasmic and is not present in tri-snRNP (Gottschalk *et al*. 2001; Boon *et al*. 2007). Depletion of Aar2p does not affect pre-mRNA splicing *in vitro*, but is lethal *in vivo*, indicating that its loss prevents further rounds of splicing without inhibiting the splicing reaction itself. This species is composed of the U5 snRNA, Aar2p, Prp8p, Snu114p, and the Sm proteins. Its electron microscopic structure is remarkably similar to the 16S counterpart in humans, the 20S U5 snRNP (Kastner *et al*. 1990; Gottschalk *et al*. 2001). This cytoplasmic 16S U5 snRNP translocates to the nucleus via an NLS in Prp8p. Once in the nucleus, Brr2p replaces Aar2p (Boon *et al*. 2007). Aar2p binds the RNase H domain in the C terminus of Prp8p and sequesters the Jab1/MPN domain of Prp8p to which Brr2p binds. Some unknown nuclear kinase phosphorylates Aar2p to induce a structural change. After phosphorylation, Aar2p releases Prp8p so that Brr2p now has access to the Jab1/MPN domain. The presence of Aar2p is hypothesized regulate tri-snRNP formation by preventing the incorporation of Brr2p into the U5 snRNP (Weber *et al*. 2011). Presumably at the same time as Brr2p, other U5 proteins (Prp28p, Snu40p, and Dib1p) join to produce the other 16S U5 snRNP species lacking Aar2p (Stevens *et al*. 2001). In the reverse manner, it is hypothesized that Aar2p might replace Brr2p in the post-splicing snRNP (Weber *et al*. 2011).

1.2.7 Diseases related to pre-mRNA splicing

Due to the large number of introns in human genes and the crucial role of pre-mRNA splicing in gene expression, dysregulation of splicing can often lead to disease. In

fact, about 50% of disease-related mutations cause disruption in splicing (López-Bigas *et al*. 2005). Mutations in either the pre-mRNA or spliceosome machinery can be the underlying cause of the phenotypes (Cooper *et al*. 2009). *Cis* mutations in consensus splice site sequences are predicted to cause approximately 15% of genetic disorders (Krawczak *et al*. 1992). However, it is evident that point mutations outside of these consensus sequences affect splicing, perhaps due to changes in splicing enhancer or silencer elements (Cooper *et al*. 2009). In support of this prediction, investigations using the (cystic fibrosis transmembrane conductance regulator) CFTR gene have shown that up to 25% of silent mutations in exons still result in splicing disruptions caused by exon skipping (Pagani *et al*. 2005).

Mutations in the proteins or snRNAs that compose the spliceosome can also contribute to disease. The only known case of a spliceosomal RNA that causes disease is the mutation of U4atac in Taybi-Linder Syndrome. Three mutations found in patients are located in the 5' stem loop of U4atac and fourth in the 3' stem loop. These mutations render splicing by the minor spliceosome null, and splicing can be rescued by a compensatory mutation. This example indicates the importance of the snRNA secondary structure in function (Edery *et al*. 2011; He *et al*. 2011).

Very few mutations in the protein machinery of the spliceosome have been identified in patients, which could indicate that such perturbations in the spliceosome are generally not tolerated. One case in which the basal splicing level is affected but not lethal occurs in retinitis pigmentosa (Cooper *et al*. 2009). Mutations in the tri-snRNP components, hPrp31p, hPrp8p, and hPrp3p have all been found in retinitis pigmentosa

patients (McKie *et al*. 2001; Chakarova *et al*. 2002; Vithana *et al*. 2001). In these patients, the mutations in Prp8p cluster in its C-terminus (McKie *et al*. 2001) and interfere with its interaction with Brr2p so the levels of functional U5 snRNP, and consequently tri-snRNP, are reduced in the cell (Boon *et al*. 2007). The pronounced effect seen in the retina is thought to be due to the daily turnover of the photoreceptor components resulting in unusually high gene expression. These *prp8* mutants are hypothesized to lower the splicing capacity of cells so that a cumulative effect of reduced splicing is evident. Another possible hypothesis is that the Aar2p-U5 immature snRNP is being incorporated into spliceosomes, causing the dominant effect seen in retinitis pigmentosa (Boon *et al*. 2007).

Interestingly, the debranching enzyme Dbr1p has been linked to a nucleic acid binding protein, TDP-43, which has been found to be mutated in some neurodegenerative diseases and forms aggregates in amyotrophic lateral sclerosis (ALS), Alzheimer's and Parkinson's diseases (Hasegawa *et al*. 2007; Lagier-Tourenne *et al*. 2010; Uryu *et al*. 2008; Armakola *et al*. 2012). Recent evidence shows that the inhibition of Dbr1p function can suppress the toxicity of TDP-43. The increase in the abundance of lariat introns competes with normal RNAs for TDP-43 binding, thereby protecting cells from the negative impact of TDP-43 aggregates (Armakola *et al*. 2012). It will be intersesting to see if inhibition of hDbr1p function becomes a viable treatment for neurodegenerative diseases in general.

Additionally, HIV-1 is thought to be dependent on Dbr1p for its replication cycle. The mechanism of HIV-1 replication is similar to that of the Ty1 retrotransposon in

yeast, which has been proposed to use yDbr1p to debranch a genomic lariat RNA in the course of reverse transcription (Cheng and Menees 2004; Ye *et al*. 2005). HIV-1 replication was reduced when the levels of hDbr1p were depleted via siRNA treatments (Ye *et al*. 2005). Perhaps Dbr1p function could be targeted as a retroviral treatment as well.

Splicing dysregultion also commonly occurs in cancers. Abnormal levels of splicing regulators can have a very different influence on splicing depending on cell type (Venables *et al*. 2008). In a well understood case, overexpression of the splicing factor SF2/ASF can induce the splicing of an oncogenic form of the ribosomal protein S6K1, which regulates translation (Karni *et al*. 2007). In general, overexpression or downregulation of helicases involved in splicing have also been correlated with cancers (Abdelhaleem 2004).

1.3. RIBOSOMAL RNA PROCESSING

Saccharromyces cerevisiae produces an average of 33 ribosomes per second (Warner 1999). RNA pol I transcribes a 35S/47S transcript (yeast and human, respectively) from rDNA loci, which is rapidly organized into the 90S small subunit (SSU) processome (rev. in Henras *et al*. 2008). From this primary transcript the mature rRNAs 25/28S, 18S, and 5.8S are released through endonucleolytic and exonucleolytic cleavages. A fourth rRNA, 5S, is transcribed in a separate message by RNA Pol III (rev. in Venema and Tollervey, 1999).

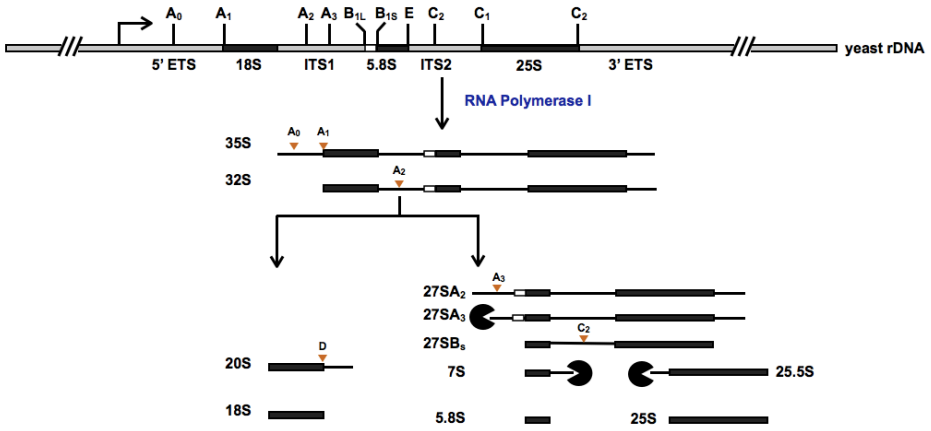
1.3.1 Organization and transcription of rDNA

Ribosomal DNA (rDNA) is organized in a cluster of 140 tandem repeats on chromosome XII in *S. cerevisiae* (Barrell *et al*. 1996). The organization of these rDNA units is conserved throughout eukaryotes, though the processing events can differ (Illustration 1.3). In yeast, the rDNA transcription units are separated by external nontranscribed spacers (ETS), which flank the 5S rRNA gene. Transcription begins at the 5' external transcribed spacer and continues through to the 3' ETS downstream of the 25S rRNA. The 18S and 25S rRNAs are separated from the 5.8S rRNA by two internal transcribed spacers (ITS1 and ITS2), which are removed through endonucleolytic and exonucleolytic cleavages (rev. in Venema and Tollervey, 1999). Co-transcriptional cleavage in the 3' ETS by Rnt1p releases the full length 35S transcript (Kufel *et al*. 1999). As Pol I transcribes the rDNA, the transcripts extending out from the rDNA begin folding elaborately and can be visualized as 'Christmas trees' by electron microscopy (Mougey *et al*. 1993). After Rnt1p cleavage in the 3' ETS, a 90S particle containing the 35S pre-rRNA transcript, the t-UTP (U Three Protein) complex, the U3 snoRNP, the UTP-B or -C complexes, and the Mpp10p complex is released (Kufel *et al*. 1999; Gallagher *et al*. 2004; Pérez-Fernández *et al*. 2007; rev. in Henras *et al*. 2008).

1.3.2 Ribosomal RNA processing pathway

The U3 snoRNP is essential for the first three cleavage events in a manner that is currently poorly understood. Without the U3 snoRNA, cleavages at the A_0, A_1, and A_2 sites are inhibited (Hughes and Ares 1991). No evidence exists that the U3 snoRNA itself

A.



B.

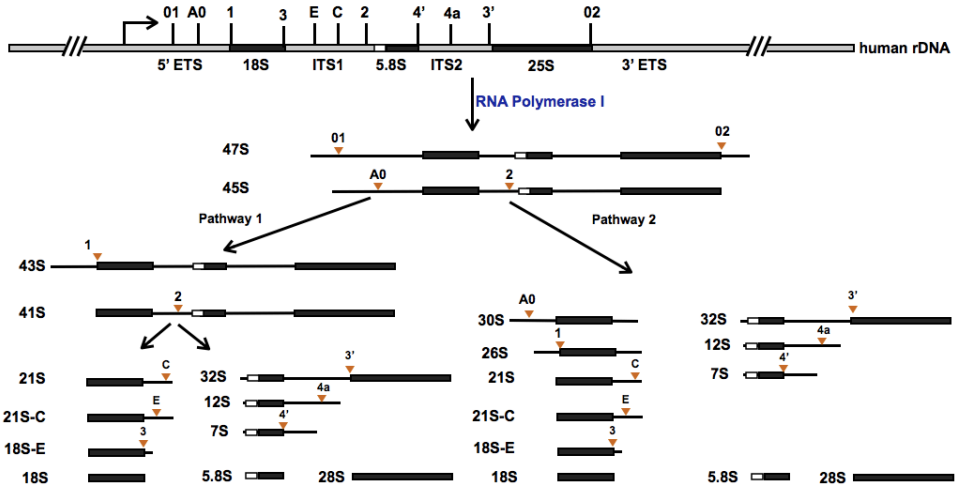


Illustration 1.3 *S. cerevisiae* and human rRNA processing pathways.

rRNA processing pathways are highly conserved from yeast (A) to humans (B) (see Chapter 1.3.2 for details). RNA Pol I initiates transcription of the rDNA at the 5' end, as indicated by the left facing arrow. Cleavage sites are indicted by the vertical lines and their associated names on the rDNA, while orange triangles indicate the cleavage events that occur at each step. The black boxes indicate mature rRNAs, while the grey boxes depict either external or internal transcribed spacers that are removed. Pacman symbols indicate exonucleolytic steps.

catalyzes the endonucleolytic reactions, and most likely the U3 snoRNP chaperones the rRNA through direct base pairing so that the 18S rRNA is properly folded (rev. in Henras *et al*. 2008). The 35S pre-RNA is first cleaved at the A_0 site in the 5' ETS, 90 nucleotides upstream of the 18S. The 33S pre-rRNA produced by cleavage at A_0 then is rapidly cleaved at the A_1 site, which immediately precedes the 18S rRNA, resulting in a rapidly processed intermediate, the 32S pre-rRNA (rev. in Venema and Tollervey, 1999). The cleavage at A_0 is not essential for cell viability and its absence does not affect A_1 cleavage; however, inhibition of cleavage at A_1 results in the loss of A_0 cleavage (Venema *et al*. 1995; rev. in Venema and Tollervey, 1999). In contrast, the A_1 and A_2 cleavages are coupled so that if A_1 cleavage is inhibited, A_2 cleavage in ITS1 does not occur. Mutants defective in A_1 cleavage produce an aberrant 22S RNA molecule that extends from A_0 to A_3, indicating that A_3 cleavage is not coupled to the previous two (Venema *et al*. 1995; rev. in Venema and Tollervey, 1999). Alternatively, an aberrant 23S product can be produced due to cleavage at A_3 prior to A_1 and A_2. The 23S RNA is not processed further, but discarded through the exosome (Zanchin and Goldfarb 1999). Rcl1p cleaves at A_2 in ITS1, splitting the 32S pre-rRNA into the 20S and $27SA_2$ pre-rRNAs (Horn *et al*. 2011). The 20S pre-rRNA is exported to the cytoplasm for further processing, while the $27SA_2$ pre-rRNA can be processed by two different pathways in the nucleus in yeast.

In the major 27S processing pathway, RNase MRP, composed of an RNA molecule and nine proteins (Chamberlain *et al*. 1998), cleaves at A_3, which is 90 nucleotides downstream of A_2, to yield $27SA_3$. Exonucleases Rat1p and/or Xrn1p digest the 5' end to the B_{1S} site, upstream of the 5' end of 5.8S, to generate $27SB_S$ (Henry *et al*.

1994; rev. in Venema and Tollervey, 1999). A second minor pathway processes $27SA_2$ by an endonucleolytic cleavage by an unknown factor directly at B_{1L}, at the 5' start of 5.8S. This pathway produces the $27SB_L$ precursor, which is only heterologous to $27SB_S$ at its 5' end (Henry *et al*. 1994). Concurrently with 5' end processing, the 3' end of $27SB_L$ and $27SB_S$ precursors are removed by an unknown mechanism up to B_2 at the end of the mature 25S rRNA. Processing at the C_1 and C_2 sites in ITS2 separates the 7S pre-rRNAs from the mature 25S rRNA (rev. in Venema and Tollervey, 1999). These reactions result in either $7SB_L$ or $7SB_S$, which are processed by 3' to 5' exonucleases to produce the $5.8S_L$ or $5.8S_S$, and both forms are efficiently incorporated into mature, functional ribosomes (rev. in Henras *et al*. 2008).

In the nucleus, a few ribosome biogenesis factors, such as Ltv1p, Pfa1p, Rio1p and Rio2p, join the pre-40S after A_2 cleavage (rev. in Henras *et al*. 2008). In the cytoplasm, the 20S pre-rRNA is methylated at two adjacent adenines, A1779 and A1780. The methylase Dim1p catalyzes both modifications, though its essential role is most likely its presence rather than its methylase activity since methylase deficient mutants support normal pre-rRNA processing (Brand *et al*. 1977; Lafontaine *et al*. 1998). For final 18S maturation, the endonuclease Nob1p cleaves at the D-site, the junction of 18S rRNA and ITS1 (Pertschy *et al*. 2009). This action is thought to occur in a translation-like cycle, in which the pre-40S subunit joins the 60S subunit through Fun12p (eIF5b) and ITS1 lies in the mRNA channel formed between the two subunits. Fun12p functions as a GTPase and is predicted to rotate the 40S so that ITS1 is displaced from the cleft. This rotation brings Nob1p in context of the D site, which is now accessible for cleavage

(Lebaron *et al*. 2012). Once D site cleavage has occurred, Rli1p and Dom34p are necessary to dissociate the now mature 40S subunit from the 60S (Strunk *et al*. 2011). It will be interesting to see if 60S undergoes a similar quality control check in the presence of pre-40S or 40S.

1.3.3 Cytoplasmic maturation of ribosomal subunits

Quality control mechanisms check the ribosomal subunits as they assemble to ensure the proper steps are completed before the subunits are released into the translating pool. Ribosomes are exported from the nucleus in a immature state, and maturation is completed in the cytoplasm by a host of *trans*-acting factors. Often these ribosome biogenesis factors are recruited or released as steps along the subunit's cytoplasmic maturation pathway (rev. in Panse and Johnson, 2010). For example, in pre-60S subunits Jjj1 and Ssa1 work together to release a factor from the exit of the polypeptide tunnel (Lo *et al*. 2010). Additionally, factors in pre-60S subunits may act as placeholders, such as Efl1p. Efl1p has homology to elongation factor-2 (EF2), and works with the assembly factor Sdo1p to authenticate a functional P site in the maturing ribosome (Bussiere *et al*. 2012).

Similarly, in late pre-40S subunits, *trans*-acting factors mask the binding sites for translation initiation factors and the entrance of the mRNA channel. For example, the binding site for Rps10 overlaps with the export adaptor Ltv1p, and Rps26 and Pno1p, the regulator of Nob1p, also share a binding site. The presence of pre-40S assembly factors prevents premature translation initiation (Strunk *et al*. 2011). Recent work indicates that a translation-like cycle may check the functionality of both large and small subunits in an

80S-like pre-ribosomal particle during a final step in their maturation (Lebaron *et al*. 2012; Strunk *et al*. 2012). The removal or remodeling of assembly factors by DExD/H-box helicases is likely to be vital for the production of mature ribosomes.

1.3.4 Modifications of rRNA directed by snoRNAs

Pseudouridylation and 2'-O ribose methylation modifications in rRNA are catalyzed by enzymes that are directed to the site by base-pairing of either box H/ACA or box C/D small nucleolar RNAs (snoRNAs), respectively (Colau *et al*. 2004; Kiss *et al*. 2006). Some bases are directly methylated by methyltransferases like Dim1p and Bud23p (Lafontaine *et al*. 1998; White *et al*. 2008). However, such modifications, though they often cluster in functional regions of the mature ribosome, often prove to not be strictly necessary for vegetative growth (King *et al*. 2003; Liang *et al*. 2007).

Additionally, the U3, U14, snR30 and snR10 snoRNAs are essential for proper folding and the rRNA cleavages A_0, A_1, and A_2 (Henras *et al*. 2008). The U3 snoRNA is the best studied of these and makes numerous regions of basepairing with the 35S pre-rRNA transcript. The secondary structure of the U3 snoRNA can be divided into three domains: a 5' domain that basepairs to the pre-rRNA, a hinge, and the 3' C/D box motifs. The 5' domain makes contacts with the 5' ETS of the pre-rRNA and within the 18S rRNA via conserved sequences Box A, Box A' and GAC box. Additionally the 5' part of the hinge region also contacts the 5' ETS. As mentioned above, the U3 snoRNA itself does not have any known catalytic function, but rather appears to act as a chaperone. The U3:rRNA interactions have also been proposed to prevent the premature folding of the 5' end of the 18S rRNA into the conserved pseudoknot structure in the mature small subunit

(rev. in Phipps *et al*. 2011). The depletion of U3 in yeast results in the cleavage at A3 without A0-A2 producing an aberrant, non-functional 23S product which is then degraded (Hughes and Ares 1991; Morrissey and Tollervey 1993).

The U14 box C/D snoRNA base pairs in two areas near a loop in the 5' end of the 18S rRNA (Liang and Fournier 1995). The basepairing site near the D box of U14 directs the methylation of C414. On the opposite side of the loop, the other snoRNA:rRNA duplex is required for rRNA processing events (Dunbar and Baserga 1998; Liang and Fournier 2006). In the absence of U14, A_0 to A_2 cleavages do not occur, which results in rapid turnover of 20S pre-rRNA, and therefore no further 18S rRNA production (Li *et al*. 1990). The removal of U14 from pre-ribosomal particles appears to be dependent multiple RNA remodeling steps since both DEAD helicases Dbp4p, Has1p, and the DEAH-box helicase Prp43p have been shown to necessary for this action (Kos and Tollervey 2005; Liang and Fournier 2006; Bohnsack *et al*. 2009).

snR30 is an essential H/ACA box snoRNA that also affects 18S rRNA processing but has no affect on 27S pre-rRNA processing. Similar to U3 and U14 defects, loss of snR30 produces the 23S aberrant pre-rRNA product (Morrissey and Tollervey 1993). Though snR30 does not appear to direct pseudouridinylation, it still basepars with the rRNA through two essential conserved motifs in the loop that normally targets the correct uridine for pseudouridinylation. The bottom of this pseudouridine loop rather than the top binds on either side of the base of an 18S rRNA hairpin (bp 799-841), making a unique snoRNA:rRNA duplex (Fayet-Lebaron *et al*. 2009). The structure of the snR30:rRNA duplex is proposed to form for loading for 18S factors (Watkins and Bohnsack 2012).

1.3.5 Diseases related to ribosome biogenesis

Despite the fact that most ribosome biogenesis factors are essential, defects in some lead to human diseases rather than embryonic lethality. Interestingly, although a ribosomal defect would be thought to have a global affect, most diseases are tissue specific. For example, defects in two different proteins involved in early 90S processing and 18S maturation have wide effects. Defects in UTP14c/UTP14 causes male infertility due to a failure in spermatogenesis, while EMG1/EMG1mutations cause Bowen-Conradi syndrome with growth retardation and psychomotor delay (Freed *et al*. 2010). In particular, many ribosomal mutations are associated with disorders that result in hematopoietic conditions, like Diamond-Blackfan anemia, 5q-syndrome and cartilage hair hypoplasia, which manifest anemias, or Shwachman-Diamond syndrome, which causes neutropenia. Mutations in the both large and small subunit proteins have been linked to Diamond-Blackman anemia, in which the bone marrow fails to produce red blood cells (Narla and Ebert 2010). Patients have been found to carry mutations in Rps19p (Draptchinskaia *et al*. 1999), which is essential for cleavage of the ITS1 in 20S pre-RNA processing and small subunit assembly (Idol *et al*. 2007). Likewise, Rps24p mutants in Diamond-Blackman syndrome patients have been shown to be deficient for small subunit assembly and the cleavage of the 5' ETS (Choesmel *et al*. 2008). With a growing number of ribosomal mutations linked to pathologies, it is likely that the ribosomal dysfunction may be the cause (Narla and Ebert 2010). Alternatively, rare mutants can modify diseases caused by a non-ribosomal protein and intensify its manifestation (Freed *et al*. 2010).

In cancer, rRNA transcription rates are increased, as well as protein synthesis, which might be due to the loss of key checkpoints in cell cycle progression. The tumor suppressor p53 can repress Pol I transcription, and the loss of p53 in many cancers could be the cause of the upregulation of rRNA production observed in those cells (Ruggero and Pandolfi 2003). Haploinsufficiency of some r-proteins may also cause dysregulation of p53 levels. In response to nucleolar stress, r-proteins Rpl5p, Rpl11p, and Rpl23p bind the E3 ubiquitin ligase Mdm2p that targets p53 for degradation. Once bound, these r-proteins normally sequester Mdm2p to allow for p53 activation, which triggers cell death. However, in cancers, reduced levels in r-proteins result in the failure to sequester Mdm2p and p53 levels are not stabilized. Thus, haploinsufficiency of the r-proteins effectively results in the loss of the tumor suppressor p53 (Zhang and Lu 2009). Additionally, the p53 mRNA can be targeted for translation via the binding of RPL26 to its 5' or 3' UTR (Freed *et al*. 2010). Therefore, loss of function for these proteins could result in downregulation of p53. Alternatively, r-proteins are often overexpressed in cancers, which might affect translation (Zhang and Lu 2009). Altogether, ribosome biogenesis is often dysregulated in cancers and diseases and understanding the molecular mechanisms in each of these cases is essential to design treatments.

1.4 RNA HELICASES IN GENE EXPRESSION

RNA helicases are ubiquitous enzymes involved in almost all gene expression processes in cells (Jankowsky 2011). Two major superfamilies (SFs) of RNA helicases are SF1 and SF2, which descended from a common ancestral helicase. The families are present in all kingdoms of life, though the SF2 members are more prominent in gene

expression in eukaryotes. Helicases in the SF1 and SF2 families have a structurally conserved helicase core with an N-terminal P-loop ATPase domain followed by a C-terminal divergent P-loop domain (Anantharaman *et al*. 2002). These two RecA-like domains can contain at least twelve motifs that make up the core, and they are highly conserved within the SF, but only partially conserved between SFs. The first RecA-like domain encompasses motifs I-III, and the second RecA-like domain with motifs IV-VI is folded back on the first domain to create the active site. Each of these core elements serves to bind a nucleotide, generally ATP, or the substrate (Illustration 1.4). The N- and C-terminal domains flanking the helicase core are specific to each protein and are thought to provide specificity or regulation for the particular function of each protein (Jankowsky 2011).

Classically, these enzymes have been characterized as ATP-dependent helicases, but their unwinding activity may not be their primary physiological function. In fact, many of these enzymes have been shown to disrupt RNA duplexes poorly (Fairman-Williams *et al*. 2010). Perhaps a more important physiological role of RNA helicases is the ATP-dependent remodeling of RNPs. RNA helicases can displace proteins from RNA concurrent with RNA unwinding or independent of it, which indicates they may not always translocate along a duplex (Jankowsky 2011). Of the SF2 family, the largest subfamily is the DEAD box helicases. Another SF2 subfamily, DEAH-box helicases, shares a related motif II sequence to DEAD helicases; however, there are some important structural, mechanistic, and physiological differences between the two subfamilies (Fairman-Williams *et al*. 2010).

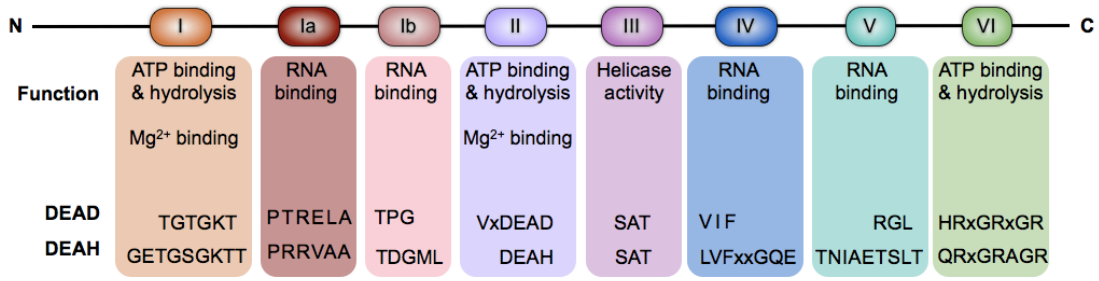


Illustration 1.4 RNA helicase core.

The core helicase motifs of the SF2 superfamily are conserved in DEAH and DEAD-box family members. Motifs I, II, and VI are involved in ATP binding and hydrolysis, while motifs Ia, Ib, IV and V all bind RNA substrates. Additionally, motifs I and II coordinate the magnesium needed for activity. The conserved sequences for the DEAD and DEAH-box families are listed beneath each motif. The characteristic name for each family is derived from the sequence of motif II.

1.4.1 DEAD-box helicases

In addition to the canonical SF2 motifs described above, the DEAD-box helicases contain the Q motif. The conserved glutamine of the Q motif specifically recognizes ATP by making direct interactions with the N6 and N7 positions of the adenosine. Therefore, DEAD-box helicases can only hydrolyze ATP (Kim *et al*. 1992; Tanner *et al*. 2003). The binding of RNA by a DEAD-box helicase is controlled by an interaction between an α-helix at the end of motif II and an arginine in motif V, which occurs only in the ATP bound state (Schütz *et al*. 2010). Interestingly, the helicase core only makes contacts along the sugar-phosphate backbone of the RNA, accounting for its lack of sequence specificity. Instead, substrate specificity is dictated by the recognition of the geometry of A form nucleic acid (Mallam *et al*. 2012) .

The DEAD-box helicases are non-processive, meaning they can locally separate duplexes. This function of DEAD-box helicases allows RNA:RNA and RNA:protein complexes to be remodeled without globally destroying the RNA structure (Cordin *et al*. 2012). DEAD-box helicases bind RNA and ATP in a cooperative manner. The N-terminal and C-terminal RecA domains function independently to bind ATP and RNA, respectively, while the helicase is in an open conformation. The interactions of each substrate binding to the opposing domains result in cooperative ATP/RNA binding, and the enzyme transitions to a closed state. Only one strand of the RNA is stably bound to domain 2, and upon closing, it is thought that domain 1 actively displaces the complimentary RNA strand. It is hypothesized that two bends are induced in the RNA by the helicase, which prevent reannealing of the strands. Also concurrent with closing, a

loop downstream of motif V moves to form the ATPase active site and possibly triggers ATP hydrolysis. With ATP hydrolysis, the dissociation of P_i and ADP allows the helicase to transition back to an open state. Thus, the enzyme is recycled (Mallam *et al*. 2012).

1.4.2 DEAH-box helicases

DEAH-box helicases contain the canonical SF2 motifs described above, as well as some distinct structural characteristics as defined by the only solved structure of a DEAH-box helicase (He *et al*. 2010; Walbott *et al*. 2010). Similar to DEAD helicases, the canonical helicase core is composed of two RecA domains with the NTP binding site between them (Fig. 4.1A). However, the nucleotide of the NTP is twisted by 150° in the DEAH-box active site, so that it is wedged non-specifically between the two RecA domains. No direct contacts are made with the nucleotide base, which accounts for the enzyme's lack of NTP specificity (He *et al*. 2010; Walbott *et al*. 2010). Therefore, DEAH-box helicases can work as NTPases, not specifically as ATPases (Schwer and Guthrie 1992). Specific insertions in the RecA domains may be a mechanism of regulation, as predicted with the human DEAH-box helicase DHX29. Structural modeling of DHX29 using Prp43p indicates a unique loop in its RecA-2 domain lies close to the putative RNA-binding tunnel. Mutational analysis proved the region to negatively regulate the basal ATPase activity of DHX29 (Dhote *et al*. 2012).

DEAH-box helicases also possess a long β-hairpin (5' HP) between motifs V and VI of the C-terminal RecA domain. In the functional model, the 5' HP moves to allow RNA binding to the active site in an ATP-dependent manner (He *et al*. 2010; Walbott *et al*. 2010). In fact, ADP binding inhibits RNA binding by the enzyme (He *et al*. 2010).

Therefore, it is thought that DEAH-box helicases can bind ATP independent of RNA binding (Cordin *et al*. 2012). This 5' HP structure is similar to the related Ski2-like helicases which possess a shorter 6-8 amino acid 5' HP; however, it appears they regulate RNA binding differently. The Ski2-like helicase Mtr4p was co-crystallized with ssRNA and ADP, indicating that ADP may not inhibit RNA binding (Weir *et al*. 2010). The 5' HP structure in both DEAH-box and Ski2-like helicases is thought to invade dsRNA to displace the duplex strands (He *et al*. 2010; Walbott *et al*. 2010; Weir *et al*. 2010). Indeed, mutations at the tip of the putative 5' HP in DHX29 abolishes its NTPase activity and renders it functionally inactive (Dhote *et al*. 2012). Also, similar to Ski2-like helicases, DEAH-box helicases contain a long α-helix known as the ratchet domain, which is proposed to pull the helicase along the RNA duplex (Büttner *et al*. 2007; He *et al*. 2010; Walbott *et al*. 2010). Unlike the non-processive DEAD-box helicases, DEAH-box helicases can use ATP hydrolysis to "ratchet" along the RNA in a reasonably processive manner (Cordin *et al*. 2012).

The DEAH-box helicases are thought to be in an open conformation when bound to ADP, and RNA binding is blocked by the 5' HP. When ATP is bound, the enzymes adopt a closed conformation, and the 5' HP is physically moved so that RNA can contact the helicase core and the ratchet domain. Upon ATP hydrolysis the enzyme is ratcheted along the duplex, and the 5' HP actively separates the two strands (He *et al*. 2010). N- or C-terminal regions may limit the movement of the two RecA domains when binding a substrate, which possibly provides some regulation of the helicase activity (He *et al*. 2010; Walbott *et al*. 2010; Weir *et al*. 2010).

1.4.3 RNA helicases in splicing

RNA helicases from the DExH/D-box family, including Ski2-like, DEAD- and DEAH-box helicases, are necessary for the proper timing and progression of the splicing cycle. DEAD helicases are needed for spliceosome assembly (Sub2p, Prp5p, and Prp28p), while the DEAH helicases (Prp2p, Prp16p, Prp22p, Prp43p) and a Ski2-like helicase (Brr2p) are required for the catalytic cycle and disassembly (King and Beggs 1990; Schwer and Guthrie 1991; Ruby *et al*. 1993; Raghunathan and Guthrie 1998b; Schwer and Gross 1998; Staley and Guthrie 1999; Kistler and Guthrie 2001; Martin *et al*. 2002; Small *et al*. 2006). While the ATPase/ helicase activity of all of these enzymes is necessary in canonical splicing, null mutants of some helicases can be bypassed by mutations in the splice site or branchpoint sequences, or by the deletion of non-essential genes with which they operate. For example, Prp5p remodels the U2 snRNA so that it can access the branchpoint region of the pre-mRNA (Wiest *et al*. 1996). However, ATPase mutants of Prp5p can be suppressed by point mutations in the branchpoint region of the intron that hyperstabilize its interaction with U2 (Xu and Query 2007). Similar results have been found for Prp16p, mutants of which can be suppressed by branchpoint mutations, and mutants of Prp22p, which can be suppressed by 3' SS mutations (Villa and Guthrie 2005; Mayas *et al*. 2006). Altogether, it is thought that the progression of the splicing reaction is regulated by the helicases through kinetic proofreading.

The kinetic proofreading model postulates that the spliceosomal helicases provide fidelity through the rate of their ATPase activity. At each step they determine whether the substrates continue through the cycle or are discarded depending on whether the ATPase

rate is normal or abnormal, respectively (Burgess and Guthrie, 1993; Illustration 1.5). During spliceosome assembly, Prp5p proofreads the U2/intron branch-site duplex (Xu and Query 2007) and Prp28p proofreads the 5' SS:U6 duplex (Yang *et al*. 2013). During catalysis, Prp16p proofreads the branch site and 5' SS before the first step (Koodathingal *et al*. 2010; Schwer and Guthrie 1992), and Prp22p proofreads the exon ligation before the second step (Mayas *et al*. 2006). Prp43p can discard stalled complexes at either the first or second steps; therefore, it is a general disassembly factor and deregulation of its helicase activity could result in inappropriate pre-mRNA discard (Koodathingal *et al*. 2010).

Additionally, these helicases can remove spliceosomal factors in a temporal fashion to contribute to fidelity. For example, the ATPase activity of Prp5p removes Cus2p in the association of U2 with the branchpoint during spliceosome assembly. In strains lacking Cus2p, this step is now ATP-independent and decoupled from Prp5p activity (Perriman *et al*. 2003). Similarly the deletion or mutation of other spliceosomal proteins can suppress DEAD-box helicase mutants, as with the case with Mud2p and helicase Sub2p, (Kistler and Guthrie 2001) or Yhc1p and helicase Prp28p (Chen *et al*. 2001). In the spliceosome, the substrates of DEAD-box helicases appear to the inherently unstable. These substrates are temporarily stabilized by other proteins, which are then displaced due to helicase activity. Without these stabilizing proteins, the ATP-dependent function of the DEAD-box helicases can be surpassed. Such data indicate that the ATP-dependent function of these helicases is truly the timing of the rearrangement reactions rather than forcing these rearrangements (Cordin *et al*. 2012).

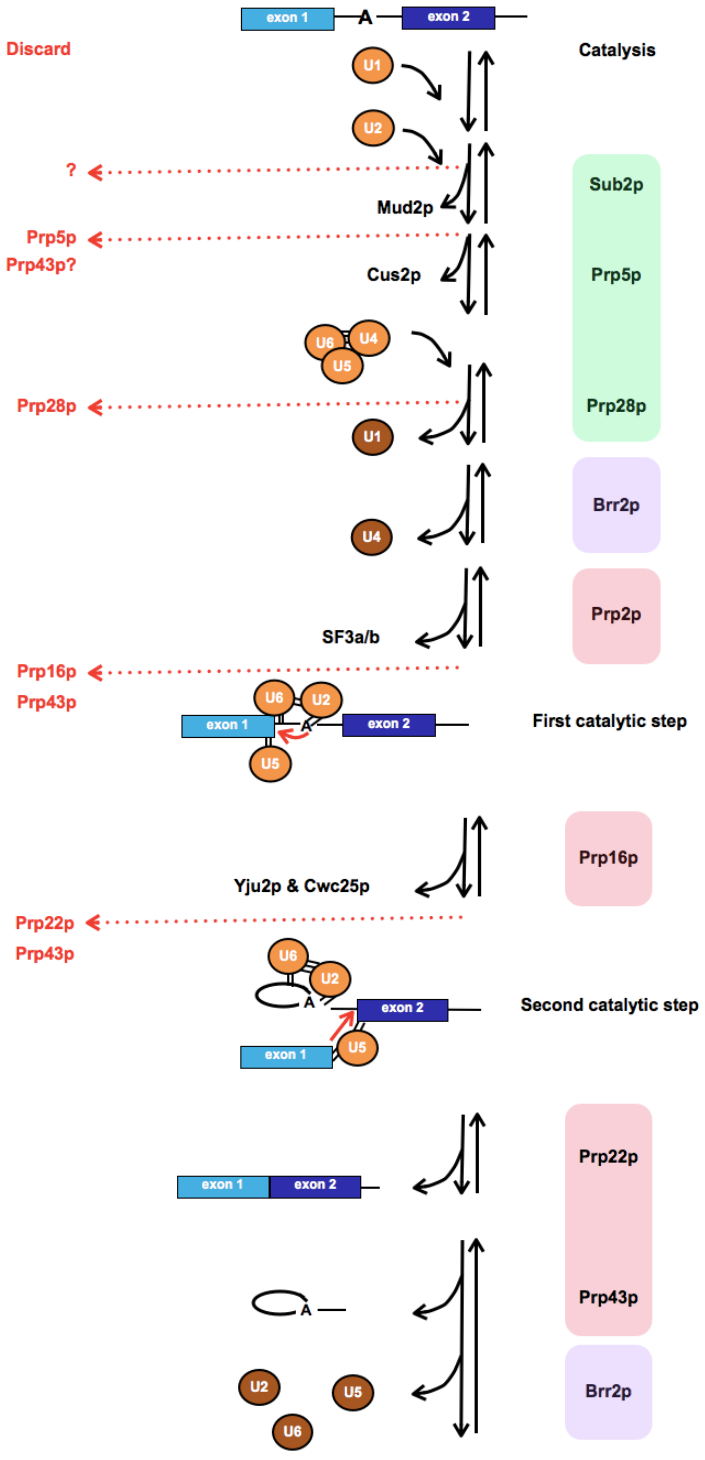


Illustration 1.5 Kinetic proofreading of pre-mRNA splicing.

(Illustration 1.5 contd.)
The splicing reaction is regulated by the ATPase activities of DEAD and DEAH-box helicases. DEAD- and DEAH-box helicases are highlighted in green and red, respectively. Brr2p, the Ski2-like helicase, is boxed in purple. The snRNPs (orange) join the pre-mRNA for the splicing reaction and are released to be recycled (dark orange). The steps at which the helicases can trigger the discard of aberrant spliceosomes are indicated by the dashed red arrows. The kinetic proofreading capability of Prp16p and Prp22p triggers the discard either before the first or second catalytic step. Prp43p is able fully discard the stalled spliceosome intermediates. Adapted from Cordin *et al*. 2012.

The RNA helicases may also facilitate optimal splicing by acting as cofactors to one another. As one helicase joins the spliceosome, it acts and undergoes a conformational change. This change then allows this helicase to function as an effector to a downstream helicase to stimulate its activity. Such a coordinated scheme would allow helicases to provide directionality to the splicing reaction (Van Nues and Beggs, 2001; rev. in Cordin *et al*. 2012).

1.4.4 RNA helicases in ribosome biogenesis

In ribosome biogenesis, RNA helicases are proposed to act in one or more of the following ways: 1) remodel RNPs locally and disrupt RNA:RNA, RNA:protein or protein:protein interactions, 2) expose pre-rRNA for nucleases or mask pre-rRNA from them, and/or 3) unwind snoRNA:pre-rRNA duplexes to release snoRNP complexes. Through these functions, the helicases are hypothesized to provide directionality to the assembly process in a similar manner to the kinetic proofreading seen for pre-mRNA splicing (Staley and Woolford 2009; Kressler *et al*. 2010). Seven DExH/D helicases have been shown to be necessary for 40S biogenesis (Dbp4p, Dbp8p, Dhr1p, Dhr2p, Fal2p, Rok1p, and Rrp3p), and ten are required for 60S biogenesis (Dbp2p, Dbp3p, Dbp6p, Dbp7p, Dbp9p, Dbp10p, Drs1p, Mak5p, Mtr4p, and Spb4p). Only two helicases, Has1p and Prp43p, have been determined to impact both large and small subunit biogenesis (rev. in Kressler *et al*. 2010). While mutations in these helicases often result in varying ribosome biogenesis defects (Bernstein *et al*. 2006; Granneman *et al*. 2006), the exact function of most of these enzymes remains elusive.

Similar to Prp43p, Has1p is necessary for cleavages at A_0, A_1, and A_2 (Emery *et al*. 2004; Combs *et al*. 2006). Loss of its function results in an increase in the 35S pre-rRNA transcript, with a stark reduction in 18S and a delay in 25S production. Depletion of Has1p results in increased amounts of snoRNAs stably associating with pre-ribosomal complexes. It should be noted that these snoRNAs bind in drastically different locations on the pre-rRNA and are of different snoRNA classes. For example, U3, U14, snR30 and snR10 are affected and are all involved in the processing of the rRNA, while snR63, snR70 and snR77 are all C/D snoRNAs. U14 and snR70 bind to the 18S rRNA, but snR63, snR77 and snR10 bind to the 25S rRNA (Emery *et al*. 2004; Liang and Fournier 2006). Interestingly, two other DEAD-box helicases have been reported to affect the association of diverse snoRNAs, Dbp4p and Rok1p (Kos and Tollervey 2005; Bohnsack *et al*. 2008). Dbp4p is also involved in the release of U14, and loss of its ATPase activity traps the snoRNA on pre-ribosomal particles (Kos and Tollervey 2005).

1.4.5 Cofactors and RNA helicases

The activity of some DExD/H-box helicases has been shown to be modulated by interacting proteins (Lebaron *et al*. 2009; Maeder *et al*. 2009; Pertschy *et al*. 2009; Montpetit *et al*. 2011; Noble *et al*. 2011). For example, the C-terminal portion (CTD) of the core spliceosomal factor Prp8p inhibits the ATPase activity of the Ski2-like helicase Brr2p, but also stimulates the helicase activity of Brr2p. This interaction between Brr2p and the CTD of Prp8p is hypothesized to account for the correct timing of U4/U6 unwinding by Brr2p in spliceosome activation (Maeder *et al*. 2009). Alternatively, some helicases can act as effector proteins for other helicases along a gene expression pathway.

Conformational changes induced by one helicase can create the substrate for the next helicase in sequence thereby providing directionality to the splicing reaction and correct temporal control (Cordin *et al*. 2012).

Regulation may be provided in DEAH-box and Ski2-like helicases through their C-terminal domains, which are involved in RNA binding and interact with the RecA domains. These domains show sequence and structural conservation and may be important in regulating the activity of the RNA helicases. Effector proteins that bind to C-terminal domains might cause the active site to be restructured so that ATP hydrolysis or helicases activity is activated or inhibited, as required. In contrast, the C-terminal domains of DEAD-box helicases are diverse. Each protein contains different features, like RRMs or dimerization motifs, that may provide specific regulation (Cordin *et al*. 2012). Additionally, DEAD-box helicases have a conserved surface of their RecA-2 domains which can bind effectors (Montpetit *et al*. 2011). Interestingly, there are three known DEAD-box box and SF1 regulators which all show structural homology. Gle1p and eIF4G have been shown to regulate Dbp5p in mRNA export and eIF4A in translation initiation, respectively (Oberer *et al*. 2005; Schütz *et al*. 2008; Montpetit *et al*. 2011). This homology suggests there may be a common activation mechanism utilized by these Gle1p- and eiF4G-like cofactors. The activators stably interact with the C-terminal RecA domain and transiently interact with the N-terminal RecA domain. In the case of Dbp5p, when it binds RNA, a closed conformation is formed. In this closed state, ATP is bound and hydrolyzed. Subsequently, Gle1p-Inositol-6-phosphate (Gle1p-IP6) binds and restructures the active site so that the RNA substrate is released. Another releasing factor,

Nup159, then binds Dbp5p to promote an open conformation, releasing ADP and Gle1p-IP6. Both of the factors are hypothesized to influence the position of the N terminal RecA domain of Dbp5p relative to its C terminal one (Montpetit *et al*. 2011).

Cofactors may also provide regulation by acting as recruitment factors (Cordin *et al*. 2012). One class of RNA binding proteins that have been implicated in providing regulation for helicases are G-patch proteins, which contain a glycine rich domain from which their name is derived. The G-patch sequence contains six glycines that follow two bulky hydrophobic residues and are separated by an aromatic residue or any amino acid (hhx(3)Gax(2)GxGhGx(4)Gx(13-20)G; h=bulky hydrophobic, a=aromatic, x=any residue for the number of positions indicated). The secondary structure of this sequence is predicted to be two α-helices connected a loop in which most of the glycines lie, and the position of the final conserved glycine is specific to each protein. Interestingly, G-patch proteins only contain one G-patch motif, and they commonly also contain other RNA binding motifs (Aravind and Koonin 1999).

There are five G-patch proteins annotated in *S. cerevisiae*: Spp2p, Ntr1p, Gno1p, Pfa1p, and YLR217W (Aravind and Koonin 1999; Silverman *et al*. 2004). Two G-patch proteins are known to recruit two spliceosomal DEAH-box helicases in yeast. Prp2p joins the spliceosome through an interaction between its C-terminus and the G-patch domain of Spp2p. An eleven amino acid sequence in Prp2p is required for Spp2p interaction, and therefore, splicing activity (Silverman *et al*. 2004). However, it is still unknown whether Spp2p can stimulate the activity of Prp2p as well as recruit it. Prp43p is also recruited by the G-patch protein Ntr1p (see Chapter 3.1 for details) (Tsai *et al*. 2007). There are no

obvious G-patch proteins known to recruit or activate Prp16p or Prp22p in yeast; however, recent data identifies GPKOW as a G-patch protein that might regulate these helicases in humans (hPrp2 and hPrp16p) (Hegele *et al*. 2012). GPKOW is necessary for splicing *in vitro*, though it does not appear to be recruiting the helicases. Mostly likely its G-patch motif is necessary for stimulating hPrp2p, making it a potential human ortholog of Spp2p (Lin RJ, personal communication). Other G-patch proteins have been identified in humans that also have potentially interesting interactions with helicases (see Chapter 5.2 for details).

1.4.6 Prp43 function in gene expression

Of the DEAH-box helicases, Prp43p is one of two that is known to be involved in more than one cellular pathway, ribosome biogenesis and pre-mRNA splicing (Jankowsky 2011). In *S. cerevisiae* Prp43p is essential and can be functionally complemented by its homolog in higher eukaryotes (Arenas and Abelson 1997; Gee *et al*. 1997). Its ATPase activity is essential for cell viability, as evidenced by lethality in *prp43* mutants harboring alanine substitutions in the conserved ATPase motifs (Martin *et al*. 2002). Prp43p uses its ATPase activity to preform different functions in the two pathways in which it is associated (Martin *et al*. 2002; Combs *et al*. 2006).

1.4.6.1 Prp43p function in pre-mRNA splicing

Prp43p was first characterized as a splicing factor whose ATPase activity is necessary for cell survival. Temperature sensitive mutants of *PRP43* were created based on mutations in other spliceosomal helicases and found to accumulate both actin pre-mRNA and excised lariat introns, indicating that Prp43p acts at the end of splicing in

spliceosome disassembly (Arenas and Abelson 1997). The cold sensitive *prp43Q423N* mutant accumulates both *in vivo* actin and U3A snoRNA introns (Leeds *et al*. 2006); however, it is not specified whether this is the lariat or as pre-mRNA. In *in vitro* splicing assays, the dominant negative mutant prp43T123Ap, which has a 94% reduced ATPase activity, accumulates the lariat intron retained within the post-splicing lariat RNP. Therefore, Prp43p promotes spliceosome disassembly by removing the lariat intron from the post-splicing lariat RNP. The released lariat intron is then debranched by Dbr1p. Also *in vitro* assays indicate that the T123A mutant does not appear impede the production of mature mRNA (Martin *et al*. 2002). In addition to its established role in canonical disassembly, Prp43p has been shown to disassemble stalled spliceosome intermediates either at the first or second step of splicing, making Prp43p a general disassembly factor. Prp43p prevents the accumulation of stalled spliceosomes bound to suboptimal substrates by discarding the substrates at steps throughout the splicing cycle (Mayas *et al*. 2006). In fact, it can act even before the 5' SS cleavage event (Koodathingal *et al*. 2010).

Prp43p is stably associated with the lariat intron and U2, U5, and U6 snRNAs, which are a part of the corresponding snRNPs in the post-splicing lariat RNP. It does not associate stably with either U1 or U4 snRNAs (Combs *et al*. 2006). However, Prp43p only crosslinks to U6 snRNA, but not the U5 or U2 snRNAs. Prp43p binding to U6 was mapped to positions 18-43 and 76-83, with this second peak partially overlapping with the U4 binding site in U6 (Bohnsack *et al*. 2009).

Ntr1 and Ntr2 proteins, which have a weak association to the NTC (Prp19-associated complex, see 1.2.4 for details), are important for the association of Prp43p

with the spliceosome. The Ntr1p/Ntr2p/Prp43p (NTR) complex can disassemble the spliceosome in an ATP-dependent manner, so that the snRNPs, lariat intron, and NTC are separated (Tsai *et al*. 2005). Ntr1p/Ntr2p can bind the spliceosome in the absence of Prp43p, and then recruit it to form the NTR complex on the spliceosome. While the addition of Prp43p or Ntr1p/Ntr2p is not an ATP-dependent step, ATP is needed for disassembly. Ntr2p mediates the association between the NTR complex and Brr2p as a component of the U5 snRNP (see Chapter 3.1 for details) (Tsai *et al*. 2007).

1.4.6.2 Prp43p function in ribosome biogenesis

Microarray analyses using recessive *prp43* ATPase mutants indicated a strong reduction in ribosomal RNAs (18S, 25S, 5.8S and 5S). Additionally, these gene expression profiles cluster *PRP43* with other ribosome biogenesis genes rather than splicing genes (Combs *et al*. 2006). Consistent with these data, Prp43p was shown to stably associate with snoRNAs and rRNA precursors. Immunoprecipitation of Prp43p indicated it associates with the 35S primary pre-rRNA transcript, as well as both the 27S and 20S pre-rRNA transcripts (Combs *et al*. 2006; Leeds *et al*. 2006; Lebaron *et al*. 2005). In contrast, Prp43p does not associate with the mature rRNAs (Combs *et al*. 2006). *prp43* ATPase mutants result in an early ribosomal processing defect, so that the 35S pre-rRNA accumulates and all other downstream precursors are reduced (Leeds *et al*. 2006; Combs *et al*. 2006). These defects in rRNA processing are evident in polysome profiles as well. *prp43G429A* and *prp43S247A* mutants have reduced levels of 80S ribosomes and the presence of halfmers was noted, indicating an imbalance in the ratio of subunits. These mutants do not properly export the precursor ribosomes; both Rpl25p and

Rps2p, large and small subunit factors respectively, accumulate in the nucleoli (Combs *et al*. 2006).

Prp43p also interacts with both box C/D (e.g. U14, snR64 and snR47) and box H/ACA (e.g. snR46 and snR3) snoRNAs (Combs *et al*. 2006). Loss of Prp43p function disrupts the association or removal of snoRNAs from the rRNA, which can result in loss of methylation (Bohnsack *et al*. 2009; Leeds *et al*. 2006). *prp43* ATPase mutants also result in increased levels of snR41, snR39, snR50, snR39b, and snR59 remaining on the ribosome. In contrast, reduced levels of snR64 and snR65 are associated with the ribosome (Bohnsack *et al*. 2009). The reduction in snR64 association is correlated to the loss of 2'-O-methylation at C2337 in the 25S rRNA (Leeds *et al*. 2006). Though cross-linking and cDNA analysis (CRAC) data does not indicate that Prp43p binds near the C2337 site, the hypothesis is that its action remodeling the rRNA elsewhere could allow this site to be exposed to allow the snoRNA to bind. CRAC data indicate that multiple Prp43p binding sites exist along the rRNA. CRAC data indicate that Prp43p binds near helix 44 at the 3' end of the 18S rRNA (ntds. 1713-1727 and 1743-1762) (Fig. 1.1). This helix is located near the D site (ntd. 1800) and the dimethylation site of Dim1p (ntd. 1781 & 1782). While this is the only predicted binding site within the 18S rRNA, there are four predicted binding sites within the 25S rRNA: bridging helix 23 & 24 (ntds. 360-378), between helix 39 & 40 (ntds. 1141-1154), along helix 34 (ntds. 703-717), and along helix 84 (ntds. 2656-2675) (Fig. 1.2). Nearby helix 34 of the LSU, the snoRNAs snR39b and snR39 or snR59 methylate positions G805 and A809, respectively. Prp43p appears to be necessary in removing these snoRNAs from the ribosome, since loss of Prp43p function

results in increased levels of these snoRNAs on the ribosome. Additionally, snR50 also methylates nearby at position G867 and is affected similarly in the absence of Prp43p function. In contrast, decreased levels of snR67, which methylates at U2724 near helix 83, were observed on the ribosomal precursors under these conditions (Bohnsack *et al*. 2009). Overall, Prp43p has been shown to extensively remodel the pre-rRNA during the course of maturation. The evidence for the action of Prp43p to allow the D site cleavage event is strengthened by its genetic interaction with Ltv1p, a non-essential component of the late pre-40S (see Chapter 2.1) (Schäfer *et al*. 2006; Pertschy *et al*. 2009).

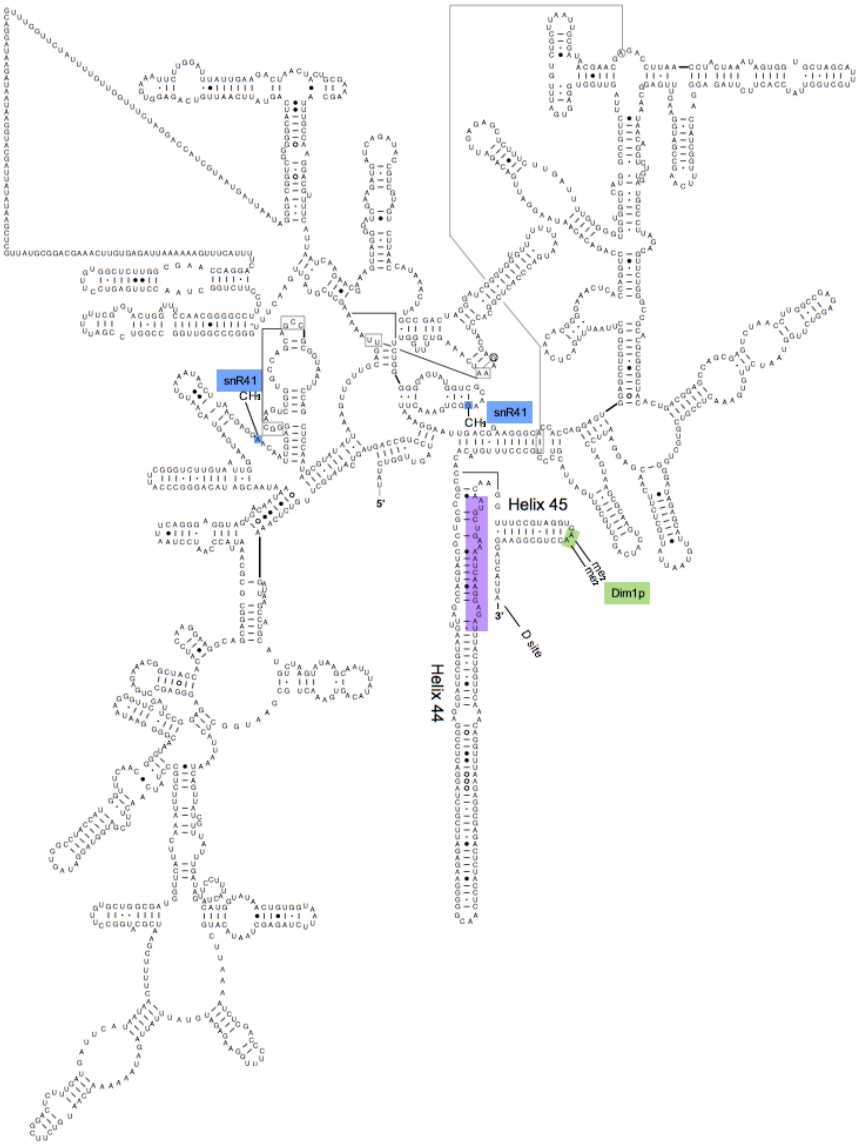


Figure 1.1 Prp43p binding site on 18S *S. cerevisiae* rRNA.

Prp43p binding sites mapped to the secondary structure of the 18S rRNA to helix 44 (in purple). The site is shown in spatial relationship to the Dim1p methylation sites (in green) and the methylation sites directed by snR41 (in blue). The D cleavage site at nucleotide 1800 also lies near the Prp43p binding site (as indicated) The secondary structure of the yeast 18S rRNA was obtained from the CRW project (http://www.rna.ccbb.utexas.edu/) (Cannone *et al*. 2002) .

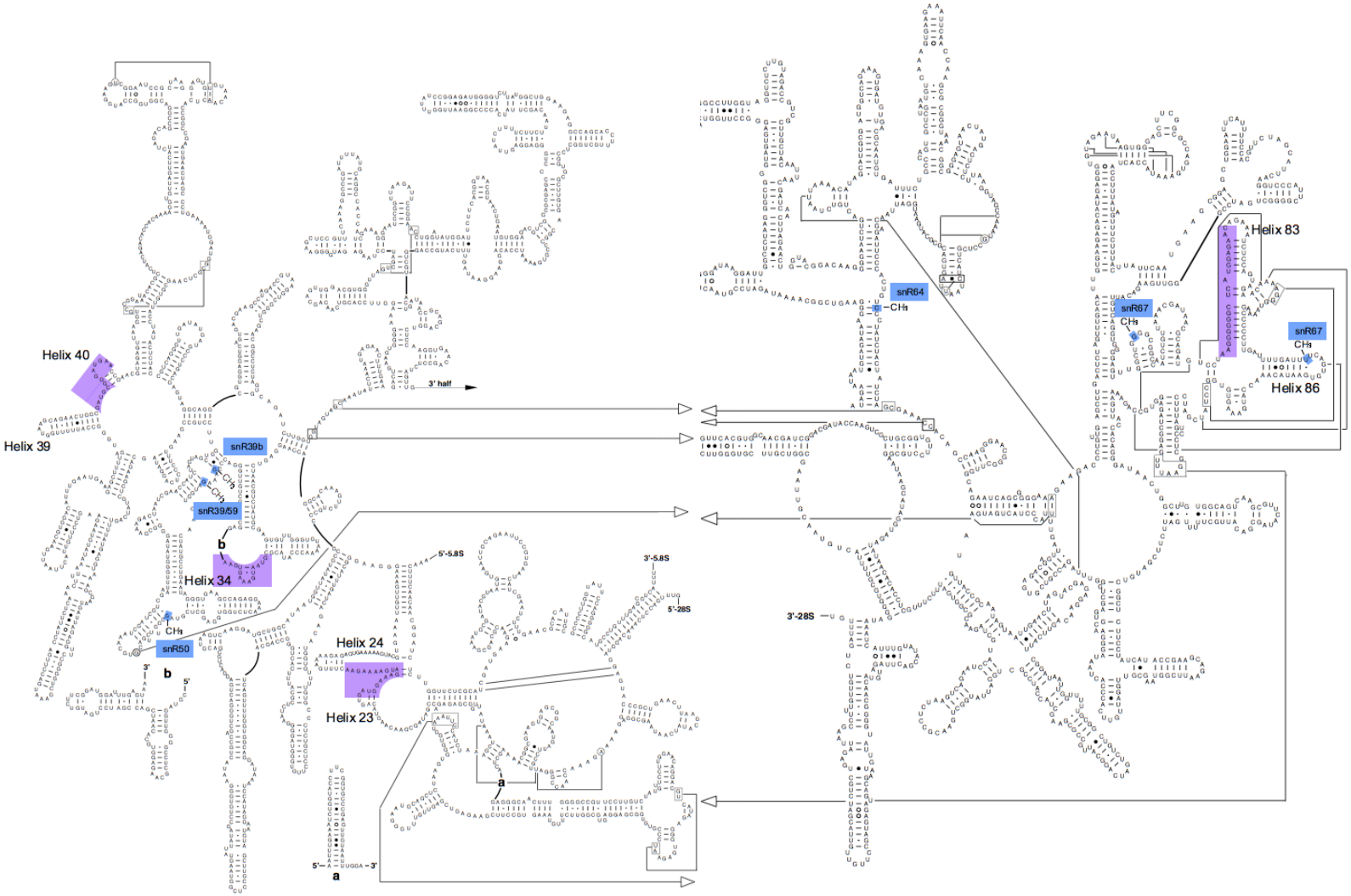


Figure 1.2 Prp43p binding sites on 25S *S. cerevisiae* rRNA.

(Figure 1.2 contd.)

Prp43p binding sites map on the secondary structure of the 25S rRNA to multiple places (purple). The bindings sites of the snoRNAs that are affected by Prp43p action are highlighted in blue. Sites for the snoRNAs snR39/snR59 and sNR39b lie in close proximity to the binding site of Prp43p on helix 34. Additionally, both modification sites of snR67 are shown in close relationship to the binding site of Prp43p on helix 83. The secondary structure of yeast 18S rRNA was obtained from the CRW project (http://www.rna.ccbb.utexas.edu/) (Cannone *et al*. 2002).

1.5 DISSERTATION OBJECTIVES

Our lab has previously proposed that Prp43p works as a communication conduit that links RNA Pol I and RNA Pol II processing pathways (Combs *et al*. 2006). If Prp43p plays a role in the crosstalk between the two pathways, we would expect its activity to be closely regulated by cells. Moreover, Prp43p localization within cells might be controlled since these two processes are physically separated in different nuclear subcompartments, with ribosome biogenesis occurring in the nucleolus and pre-mRNA splicing occurring in the nucleoplasm (rev. in Wahl *et al*. 2009). Shortly after beginning my work, the G-patch protein Ntr1p was found to recruit Prp43p for spliceosomal disassembly (Tsai *et al*. 2007; Tanaka *et al*. 2007). It was unknown how Prp43p is directed to its substrates in ribosome biogenesis, but the interaction of Prp43p with two other G-patch proteins suggested they might be similarly involved in Prp43p recruitment. My dissertation work explored the possibility that these three G-patch proteins work together to modulate Prp43p within cells. I then set out to determine if and how the G-patch proteins related to Prp43p affect ribosome biogenesis and pre-mRNA splicing. I hypothesized that each G-patch protein interacts with Prp43p in a mutually exclusive manner and competes with one another to recruit Prp43p to spliceosome disassembly, early or late ribosome biogenesis. In this aim, I sought to understand if and how the G-patch proteins impact one another and their relationships with Prp43p. From my data, I built a network of G-patch protein and Prp43p interactions.

Additionally, Prp43p does not have inherent substrate specificity (Tanaka *et al*. 2007), which suggests it must be directed to its substrates and/or stimulated within the

appropriate ribonucleoprotein context. Since the C-terminus of Prp43p is unique, I set out to create and characterize *prp43* alleles with mutations in its C-terminus that only affect either pre-mRNA splicing or ribosome biogenesis. By characterizing these alleles we hoped to determine specificity domains of Prp43p, which we predicted to lie in its C-terminus. These domains might be interaction surfaces for the G-patch proteins or innate regions with differential functions. The work was begun before the crystal structure of Prp43p was published (He *et al*. 2010, Walbott *et al*. 2010), but we feel that the structure provides crucial insight into the nature of the alleles I have identified.

My findings contribute to our understanding of the relationship between Prp43p and its cofactors and as well as provide *in vivo* data for *prp43* mutants with a range of effects on ribosome biogenesis and splicing. I believe that my work will be useful in building a cohesive model for Prp43p regulation and function. Since Prp43p is conserved in higher eukaryotes we anticipate our understanding of Prp43p function in *S. cerevisiae* will be instructive in unraveling its function and regulation in humans.

Chapter 2: Pfa1p acts as a negative regulator of splicing and early ribosome biogenesis by modulating Prp43p trafficking

2.1 Introduction to Pfa1p

Pfa1p (Prp Forty-three Associated) was first identified as an interacting factor with Prp43p by mass spectrometry and it proved to be the most abundant protein present in Prp43p purified material. Pfa1p was implicated as a component of pre-40S particles due to its association with the 20S pre-rRNA, and possibly 18S rRNA. It does not associate with the 35S or 27S pre-rRNA species, nor the mature 25S rRNA (Lebaron *et al*. 2005). Pfa1p was further shown to be partially responsible for the maturation of the pre-40S subunit (Pertschy *et al*. 2009), and deletion of *PFA1* reduces the efficiency of Prp43p association with 20S pre-rRNA (Lebaron *et al*. 2009). The data suggest that Pfa1p also interacts with Prp43p outside the context of the pre-40S since Prp43p appears to be stoichiometric with Pfa1p and in excess over the co-purified pre-40S components in Pfa1p-TAP purified material (Lebaron *et al*. 2009).

2.1.1 Role of Pfa1p in small subunit ribosome biogenesis

Pfa1p was implicated in the maturation of the small subunit through a negative genetic interaction with the small subunit factor *LTV1* (Pertschy *et al*. 2009). Ltv1p was characterized as a late pre-40S component associated with Enp1, Hrr25p, Nob1p, Rio1p, Tsr1p, Dim1p and Dim2p (Schäfer *et al*. 2006). Ltv1p has also been proposed to function as an NES for pre-40S subunits in a Crm1-dependent manner; however, export is only partially blocked in an *ltv1Δ* strain (Seiser *et al*. 2006). Strain background appears to

affect the extent of the *ltv1Δ* on phenotype; in a W303 background *ltv1Δ* has minimal growth defects at 23°C and 30°C (Pertschy *et al*. 2009), but in the S288C background the *ltv1Δ* mutant is significantly cold sensitive (Seiser *et al*. 2006). The growth defects observed in *ltv1Δ* at 23°C are exacerbated in the *ltv1Δpfa1Δ* double mutant. Interestingly, the loss of the G-patch domain of Pfa1p is sufficient to produce this effect, indicating this region of the protein is responsible for this genetic interaction (Pertschy *et al*. 2009). In an *ltv1Δ* mutant the processing of the 20S pre-rRNA into the 18S rRNA is aberrant, and the 20S pre-rRNA partially accumulates in the cytoplasm (Seiser *et al*. 2006; Pertschy *et al*. 2009). The *ltv1Δpfa1Δ* double mutant not only accumulates 20S pre-rRNA, but an unusual 17S species also accumulates due to the incorrect trimming of the 20S pre-rRNA by the exosome. The accumulation of the 20S pre-rRNA in the cytoplasm is greatly exacerbated in the *ltv1Δpfa1Δ* double mutant compared to the *ltv1Δ*. Interestingly, the growth defects and 20S processing defects in the *ltv1Δpfa1Δ* double mutant can be rescued by overexpression of Nob1p. The endonuclease Nob1p cleaves at the D site of 20S pre-rRNA to yield the 18S rRNA. Nob1p mutants that are deficient in the endonuclease activity are unable to rescue the phenotypes observed in *ltv1Δpfa1Δ*, indicating its activity is necessary for the suppression. While *ltv1Δpfa1Δ* has ribosome biogenesis defects, *pfa1Δ* alone does not. The loss of Pfa1p does not result in rRNA processing defects or growth defects at any temperature. Ltv1p also genetically interacts with an allele of *PRP43*, *prp43S414F* (Pertschy *et al*. 2009). This point mutation lies near the base of the β hairpin between motifs V and VI (Fig. 4.1A). Despite a mutation in such a critical area of the helicase, *prp43S414F* does not have any obvious growth or rRNA

processing defects. However, the *ltv1Δ prp43S414F* double mutant has 20S pre-rRNA processing defects similar to *ltv1Δpfa1Δ*. Similarly, Nob1p overexpression can rescue *ltv1Δprp43S414F* phenotypes. These genetic interactions lead to a model that defines Prp43p and Pfa1p as factors involved in the D site cleavage event. Together they contribute to the stability of the 3' end of the 18S rRNA and prevent exosome access (Pertschy *et al*. 2009).

Pfa1p can act as a cofactor for Prp43p by stimulating its helicase and ATPase activity. Pfa1p physically interacts with Prp43p through two structural domains at the N- and C-termini of Pfa1p (1-202 aa and 574-767 aa). Each domain can interact with Prp43p independent of the other, and the interaction is further stabilized through the presence of the intervening domain. The G-patch domain of Pfa1p lies in its C-terminus and interacts with C-terminus of Prp43p (1-657 aa) (Lebaron *et al*. 2009; Walbott *et al*. 2010). This interaction is necessary to stimulate the ATPase and helicase activity of Prp43p, and the stimulation is not based on an increased binding of ATP by Prp43p (Lebaron *et al*. 2009). The model for the role of Prp43p in small subunit biogenesis postulates that the stimulation of Prp43p by Pfa1p is needed to restructure the rRNA at helix 44 (Fig. 1.1) to allow Nob1p access to the D site (Pertschy *et al*. 2009; Lebaron *et al*. 2009). Nob1p is prevented from acting prematurely in late pre-40S particles, either by proteins or RNA duplexes which are resolved by Prp43p/Pfap1p (Pertschy *et al*. 2009; Lebaron *et al*. 2009; Walbott *et al*. 2010). Alternatively, the effect of Prp43p and Pfa1p on 20S pre-rRNA processing could be secondary. They could be affecting the overall structure of the pre-rRNA due to an inability to remove snoRNAs or proteins. Since neither *pfa1Δ* nor

prp43S414F mutants have any growth defects, the essentiality of this regulation is unclear. Further work will need to done to understand if the role if Prp43p in 20S processing is direct or indirect.

2.1.2 Effect of Pfa1p on splicing efficiency

In contrast to the *pfa1Δ* mutant, overexpression of Pfa1p results in a slow growth effect (Pandit *et al*. 2009). The deletion of the G-patch domain from Pfa1p does not alleviate its effect on growth when overexpressed; therefore, regions of Pfa1p outside the G-patch domain drive the growth defect (Pandit *et al*. 2009). Since the G-patch domain of Pfa1p is necessary for stimulating the ATPase and helicase activity of Prp43p, (Lebaron *et al*. 2009), the overstimulation of Prp43p activity is not likely to the underlying cause of the defect.

Excess Pfa1p results in a first step pre-mRNA splicing defect so that pre-mRNA accumulates (Pandit *et al*. 2009). However, it is unlikely that Pfa1p functions directly in splicing as it does not stably interact with any snRNAs (Lebaron *et al*. 2005). Increased levels of Pfa1p can also affect rRNA processing so that all pre-rRNAs are downregulated (Pandit *et al*. 2009). Our results show different rRNA processing defects during Pfa1p overexpressed compared to previously published data. These discrepancies are perhaps due to strain background, considering the differences seen in the *ltv1Δ* mutant in different strains. The previous studies used a W303 background (Pandit *et al*. 2006) while we have used a S288C background.

To understand how Prp43p is regulated through cofactors, we further investigated its relationship with Pfa1p. We sought to understand how Pfa1p influences splicing and

ribosome biogenesis with the hypothesis that overexpressed Pfa1p sequesters Prp43p from participating in both processes. Since Pfa1p is involved in a late ribosome biogenesis step, we considered how its overexpression might be affecting the function of Prp43p at an earlier step. Our data indicate that Pfa1p can act as a negative regulator of Prp43p, possibly functioning to pause gene expression in response to environmental stress conditions.

2.2 MATERIALS AND METHODS

2.2.1 Construction of yeast strains

TAP-tagged strains were obtained from the TAP collection and verified by Western blotting (Howson *et al*. 2005). C-terminal GFP tagging of Prp43p was achieved through PCR amplification of the eGFP tag and *KANr* (Kanamycin resistance) marker from the genomic DNA of strain ss599 (Combs *et al*. 2006). The PCR amplified epitope cassette was introduced via homologous recombination by lithium acetate transformation as previously described (Gietz and Woods 2002). Colonies growing on YPD media supplemented with 200 μg/mL geneticin (Gibco BRL Life Technologies, Inc.) were screened by Western blotting and fluorescence for proper integration of the GFP tag. The C-terminal myc tags were amplified from p*FA6a-13Myc-TRP1* (Longtine *et al*. 1998) or p*FA6a-13Myc-URA3* vectors. PCR amplification from the p*FA6a-13Myc-URA3* plasmid utilized different reverse primers from the other p*FA6a* plasmids: Pfa1R1, Gno1R1, and Ntr1R1. The N-terminal *HIS3-GAL1-3HA* cassette was PCR amplified from the p*FA6a-HIS3-GAL-3HA* vector (Longtine *et al*. 1998) using the primers listed in Table 2.2. All tags were integrated via homologous recombination as described above. MYC tagged

Table 2.1 Yeast strains used in Chapter 2

Strain	Relevant Gene(s)	Genotype	Source/Reference
599	*PRP43-GFP*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX*	(Combs *et al*. 2006)
3067	pRS316-*PRP43-TAP pfa1Δ*	*MAT* **a** *ura3-52, trp1-63, his3Δ200, leu2-1, ade2-101, lys2-801, prp43::KANMX pfa1::HIS3*, pRS316-*PRP43-TAP*	This study
3070	*NTR1-TAP*	*MAT* **a** *his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, NTR1-TAP::HIS3*	(Howson *et al*. 2005)
3071	*PFA1-TAP*	*MAT* **a** *his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, PFA1-TAP::HIS3*	(Howson *et al*. 2005)
3072	*GNO1-TAP*	*MAT* ***a*** *his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, GNO1-TAP::HIS3*	(Howson *et al*. 2005)
3073	*NTR1-TAP PRP43-GFP*	*MAT* **a** *his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, NTR1-TAP::HIS3, PRP43-GFP::KANMX*	This study
3074	*PFA1-TAP PRP43-GFP*	*MAT* **a** *his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, PFA1-TAP::HIS3, PRP43-GFP::KANMX*	This study
3075	*GNO1-TAP PRP43-GFP*	*MAT* **a** *his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, GNO1-TAP::HIS3, PRP43-GFP::KANMX*	This study
3099	*PFA1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, PFA1-13myc::URA3*	This study
3100	*NTR1-TAP PRP43-GFP PFA1-myc*	*MAT* **a** *his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, NTR1-TAP::HIS3, PRP43-GFP::KANMX, PFA1-13myc::URA3*	This study
3101	*GNO1-TAP PRP43-GFP PFA1-myc*	*MAT* **a** *his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, GNO1-TAP::HIS3, PRP43-GFP::KANMX, PFA1-13myc::URA3*	This study
3102	*GNO1-TAP PRP43-GFP NTR1-myc*	*MAT* **a** *his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, GNO1-TAP::HIS3, PRP43-GFP::KANMX, NTR1-13myc::URA3*	This study

Strain	Relevant Gene(s)	Genotype	Source/Reference
3103	*NTR1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, NTR1-13myc::URA3*	This study
3104	*PFA1-TAP PRP43-GFP GNO1-myc*	*MAT* **a** *his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, PFA1-TAP::HIS3, PRP43-GFP::KANMX, GNO1-13myc::URA3*	This study
3107	*GNO1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, GNO1-13myc::URA3*	This study
3108	*PFA1-TAP PRP43-GFP NTR1-myc*	*MAT* **a** *his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, PFA1-TAP::HIS3, PRP43-GFP::KANMX, NTR1-13myc::URA3*	This study
3109	*NTR1-TAP PRP43-GFP GNO1-myc*	*MAT* **a** *his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ ,0 NTR1-TAP::HIS3, PRP43-GFP::KANMX, GNO1-13myc::URA3*	This study
3129	pRS316-*GAL PRP43-GFP*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX*, pRS316-*GAL*	This study
3131	p*GAL-PFA1-HA/ZZ PRP43-GFP*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX*, p*GAL-PFA1-HA/ZZ-URA*	This study
3138	p*GAL-PFA1-HA/ZZ PRP43-GFP NTR1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX, NTR1-13myc::TRP1*, p*GAL-PFA1-HA/ZZ-URA*	This study
3139	pRS316-*GAL PRP43-GFP NTR1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX, NTR1-13myc::TRP1*, pRS316-*GAL*	This study
3144	*GAL::HA-NTR1 PRP43-GFP*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, HIS3::GAL::HA-NTR1, PRP43-GFP::KANMX*	This study
3145	*GAL::HA-GNO1*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, HIS3::GAL::HA-*	This study

Strain	Relevant Gene(s)	Genotype	Source/Reference
	PRP43-GFP	*GNO1, PRP43-GFP::KANMX*	
3146	*GAL::HA-PFA1 PRP43-GFP*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, HIS3::GAL::HA-PFA1, PRP43-GFP::KANMX*	This study
3148	*GAL::HA-PFA1 PRP43-GFP NTR1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, HIS3::GAL::HA-PFA1, PRP43-GFP::KANMX, NTR1-13myc::TRP1*	This study
3149	*GAL::HA-PFA1 PRP43-GFP GNO1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, HIS3::GAL::HA-PFA1, PRP43-GFP::KANMX, GNO1-13myc::TRP1*	This study
3153	*GAL::HA-NTR1 Prp43-GFP PFA1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, HIS3::GAL::HA-NTR1, PRP43-GFP::KANMX, PFA1-13myc::TRP1*	This study
3154	*GAL::HA-NTR1 PRP43-GFP GNO1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, HIS3::GAL::HA-NTR1, PRP43-GFP::KANMX, GNO1-13myc::TRP1*	This study
3155	*GAL::HA-GNO1 PRP43-GFP NTR1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, HIS3::GAL::HA-GNO1, PRP43-GFP::KANMX, NTR1-13myc::TRP1*	This study
3156	*GAL::HA-GNO1 PRP43-GFP PFA1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, GAL:HA-GNO1::HIS3, PRP43-GFP::KANMX, PFA1-13myc::TRP1*	This study
3157	pRS316-*GAL PRP43-GFP GNO1-TAP*	*MAT **a** his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, GNO1-TAP::HIS3, PRP43-GFP:: KANMX*, pRS316-*GAL*	This study
3159	p*GAL-PFA1-HA/ZZ PRP43-GFP*	*MAT **a** his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, GNO1-TAP::HIS3, PRP43-*	This study

Strain	Relevant Gene(s)	Genotype	Source/Reference
	GNO1-TAP	*GFP::KANMX*, p*GAL-PFA1-HA/ZZ*	
3169	pRS411 *PRP43-GFP*	*MATα his3Δ1*, *leu2Δ0*, *met15Δ0*, *trp1Δ63*, *ura3Δ0*, *PRP43-GFP::KANMX*, pRS411	This study
3170	pRS411 *PRP43-GFP GAL::HA-PFA1*	*MATα his3Δ1*, *leu2Δ0*, *met15Δ0*, *trp1Δ63*, *ura3Δ0*, *PRP43-GFP::KANMX*, *HIS3::GAL::HA-PFA1*, pRS411	This study
3181	*GAL::HA-NTR1 PRP43-GFP pfa1Δ*	*MATα his3Δ1*, *leu2Δ0*, *met15Δ0*, *trp1Δ63*, *ura3Δ0*, *PRP43-GFP:: KANMX*, *HIS3::GAL::HA-NTR1*, *pfa1Δ::URA3*	This study
3182	*GAL::HA-GNO1 PRP43-GFP pfa1Δ*	*MATα his3Δ1*, *leu2Δ0*, *met15Δ0*, *trp1Δ63*, *ura3Δ0*, *PRP43-GFP:: KANMX*, *HIS3::GAL::HA-GNO1*, *pfa1Δ::URA3*	This study
3183	*PRP43-GFP pfa1Δ*	*MATα his3Δ1*, *leu2Δ0*, *met15Δ0*, *trp1Δ63*, *ura3Δ0*, *PRP43-GFP:: KANMX*, *pfa1Δ::URA3*	This study
3213	p*GAL-GNO1-HA/ZZ*	*MAT **a** his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ0*, p*GAL-GNO1-HA/ZZ*	This study
3214	p*GAL-NTR1-HA/ZZ*	*MAT **a** his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ0*, p*GAL-NTR1-HA/ZZ*	This study

Table 2.2 Oligonucleotides used in Chapter 2

Name	Sequence
Ntr1 R1	5'TCTAATCAATTTTGTTTTTCGACAATAATATATAAATCGTGCCTATCTCATCTG GAGGAAGTTTGAGAGG3'

Name	Sequence
NTR1 F2	5’AGGATTCCAGTGGGACCTTTAAGCCAATTTATTTATGGGCCCTTGACCTCGGT CGACGGATCCCCGGGTT 3’
Gno1 F2	5’CGTTGATGGACTCCAAGGCACTGAATGAGATCTTTATGATAACAAACGACGGT CGACGGATCCCCGGGT 3’
Gno1 R1	5’TCCCCACTCTATATATCTTGCTGCGTGTGCAGACTGGCCAGCTGCTCACATCTG GAGGAAGTTTGAGAGG 3’
PFA1F2	5’TTTTTGCCAAAATTAAAAAGAATAGATCGGGTTTAAGACACAGTGAAAGTGG TCGACGGATCCCCGGGTT 3’
Pfa1 R1	5’AAAATAAAAAAATATGTATATACGCAATTGTTAAGAAAATGAGGATTAAATC TGGAGGAAGTTTGAGAGG 3’
Pfa1 R1 Longtine	5’AAAATAAAAAAATATGTATATACGCAATTGTTAAGAAAATGAATTCGAGCTC GTTTAAAC3’
Ntr1 R1 Longtine	5’TTTGTTTTTCGACAATAATATATAAAATCGTGCCTATCTCAGAATTCGAGCTCG TTTAAAC 3’
Gno1 R1 Longtine	5’TATATATCTTGCTCGTGTGCAGACTGGCCAGCTGCTCACGAATTCGAGCTCGT TTAAAC 3’
E-C2	5’ GAAAAGGCCAGCAATTTCAAGTTAACTCCAAAGAGTATCACTCAC 3’
D-A2	5’ AAACTTACAAGCCTAGCAAGACCGCGCACTTAAGCGCAGGCCCGG 3’
scR1	5’ CGTGTCTAGCCGCGAGGAAGGATTTGTTCC 3’

Name	Sequence
U3 PE	5' CCAAGTTGGATTCAGTGGCTC 3'
U1	5' GAATGGAAACGTCAGCAAACA 3'
NTR1-R3	5’AAAAGAACTTTTATCTGTGTTGGAGTCCGAATCCTCCATGCACTGAGCAGCGTAATCTG 3’
PFA1-R3	5’TACGTCTGCTTCCTTGATAATGACTATGCCTTTTTGCCATGCACTGAGCAGCGTAATCTG 3’
GNO1-R3	5’AACCAAACCGCTGTTTGGTTCTTGTAGCTGCCAAACCCATGCACTGAGCAGCGTAATCTG 3’
NTR1-F4	5’TTCGAAATAGTACAACCGAGAGAGGTCGAAGAACTTAAGCGAATTCGAGCTCGTTTAAAC 3’
PFA1-F4	5’AGAAGTGATTTAATCGCAGAGTCAAATATAAGCATAAATTGAATTCGAGCTCGTTAAAC 3’
GNO1-F4	5’GAGATGAGTACTCAATAGTAACATATAGGCAGCTTACACCGAATTCGAGCTCGTTTAAAC3’
35S	5’CGCTGCTCACC AATGG3’
U1 North	5' GAATGGAAACGTCAGCAAACA 3'
U2 North	5' AAAAGAACAGATACTACACTTGA 3'

Name	Sequence
U4 North	5' ACCATGAGGAGACGGTCTGG 3'
U5 North	5' ACACCCGGAATGGTTCTGGTA 3'
U6 North	5' AACGAAATAAATCTCTTTGTAAAAC

strains were verified by PCR amplification of genomic DNA, and by western blotting, while HA tagged strains were verified by Western blotting.

2.2.2 Plasmids

The p*FA6a-13Myc-URA3* vector was a kind gift from Won-Ki Huh (Sung *et al*. 2008), but due to the lack of successful cloning with this construct the 13MYC cassette from p*FA6a-13Myc-KANr*, was subcloned into the p*FA6a-13Myc-URA* backbone using Xma1 and BglII. The following Yeast ORF clones were obtained from Open Biosystems and verified through DNA sequencing: YLR424W (YSC3869-9520126), YNL224C (YSC3869-9516572), and YGR280C (YSC3869-9514962). Yeast transformations of plasmids were performed as previously described (Gietz and Woods 2002).

2.2.3 Time Course Western Analysis

Whole cell protein extracts from indicated time points were prepared as previously described with minor modifications (Yaffe and Schatz 1984). Cell cultures were lysed after adjusting to 2N NaOH, 8% 2-mercaptoethanol. Proteins were precipitated with acetone and solubilized in 2X LDS buffer with 200 mM DTT, separated by SDS-PAGE, and transferred to nitrocellulose. Antibodies used in western blot analysis were as follows: 1:5000 anti-c-myc purified from Myc 1-9E10.2 (ATCC CRL-1729), 1:5000 anti-HA (Covance HA.11 Clone 16B12, MMS-101R), 1:5000 anti-GFP (GenScript) conjugated to HRP using Innova Biosciences Lightning-Link HRP conjugation kit, 1:5000 goat anti-mouse HRP conjugated (Rockland, 610-1302) or 1:30,000 goat anti-mouse HRP conjugated (JacksonImmunoresearch, 115-035-062), 1:10,000 Rabbit PAP (Rockland, P300-0025), 1:8000 rabbit anti-Rpl8 (A. Johnson, UT

Austin), 1:5000 anti-rabbit HRP (Rockland, 611-303-122). All blots were performed in BLOTTO [5% dry milk in 1x PBST (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, 0.2% Tween-20)] for 1 hour. Washes were performed with 1x PBST. Blots were visualized on X-ray film by chemiluminescence using Perkin Elmer Western Lightning Plus ECL substrate. Membranes used for multiple western blots were stripped with Restore western stripping buffer (Thermo Scientific) and removal of signal was determined by film or on a Molecular Dynamics Storm 860 imager

2.2.4 Serial Dilutions

Yeast strains were grown in non-repressive YP media + 2% sucrose until OD_{600} 1.0. Serial dilutions of 10^{-2} were made and 3 μL of each dilution was plated onto either YP + 2% dextrose or YP + 2% galactose plates. When selection was required, cells were grown in S-URA + 2% sucrose and plated on S-URA + 2% dextrose or 2% galactose plates. Plates were incubated at 31°C for 3 days before photographing.

2.2.5 TAP Coimmunoprecipitations

Cells were grown at 31°C in YPD, unless otherwise indicated. Cells were harvested, resuspended (1:1 w:v), frozen in AGK buffer (10 mM HEPES pH 7.9, 200 mM KCl, 1.5 mM $MgCl_2$, and 8% glycerol) and lysed cryogenically in a Retch ball mill (Mayas *et al*. 2006). Approximately 1 g of cell material was used for each of these experiments. After addition of leupeptin/pepstatin (1 mM each), DTT (1 mM), and PMSF (0.4 mM), proteins were extracted for 30 minutes by rotating at 4ºC and clarified by centrifugation at 15,000 x *g* for 20 minutes. Extracts were normalized by OD (A_{595}) using

a BioRAD Bradford assay. TAP coimmunoprecipitations were carried out using IgG sepharose resin as previously described (Puig *et al*. 2001). IgG beads were washed with IPP150 (10 mM Tris, pH 8.0, 150 mM NaCl, 0.1% NP-40, 1.5 mM $MgCl_2$, 8% glycerol, 1 mM DTT) 4X at 2500 x *g*. Bound proteins were eluted by boiling in 2X LDS + 200 mM DTT. For glycerol gradient sedimentation, IgG beads were incubated with 1 u of TEV protease (prepared in the Stevens laboratory) and 100 u of murine RNase inhibitor (NEB, M0314L) at 16˚C for 1-2 hours. The eluted material was divided and incubated with or without 2 mM ATP at 25˚C for 10 minutes before layering on a 10-30% glycerol gradient prepared in 20 mM HEPES, pH 7.9, 1.5 mM $MgCl_2$, 0.02% NP-40, 150 mM KCl, 1 mM DTT, 0.4 mM PMSF. Gradients were centrifuged at 29,000 rpm for 16 hours at 4°C in an SW-41 rotor at 4°C. Fractions were phenol/chloroform extracted, and RNA was precipitated with 0.1 volume of 3 M sodium acetate, pH 5.3 and 2.5 volumes of 100% ethanol. RNAs were separated through denaturing 8M Urea 1X TBE 40% PAGE [19:1 acrylamide:bis] gel and transferred to nylon membrane (Whatman) by electrophoresis in 25 mM sodium phosphate, pH 7.0. RNA was crosslinked to the membrane by exposure to 254 nm UV light for 1 minute before northern blotting for all snRNAs

2.2.6 GFP Coimmunoprecipitations

GFP coimmunoprecipitations were performed with either with abcam anti-GFP antibody (ab1218) or Chromotek-GFP-TRAP beads (Allele Biotech, ACT-CM-GFA0050). Extracts was prepared as described above. Using anti-GFP antibody ab1218, the extract was first incubated with rProtein G Agarose (Invitrogen, #15920-010) for 30

minutes at 4ºC with rotation. Pre-cleared extract was removed and proteins were quantitated by BioRAD Bradford assay as described above, then incubated with 1:250 v:v anti-GFP antibody at 4º C for 1 hour with rotation. Proteins were then captured on rProtein G Agarose beads for 30 minutes at 4º C with rotation. rProtein G beads were washed four times with Na_2HPO_4 IPP150 (10 mM, Na_2HPO_4 pH 7.0, 150 mM NaCl, 0.1% NP-40, 1.5 mM $MgCl_2$, 8% glycerol, 1 mM DTT) and harvested by centrifugation at 2500 x *g*. Bound material was removed either by boiling in 2X LDS + 200 mM DTT or by phenol-chloroform extraction. For experiments using Chromotek-GFP-TRAP beads, prepared extracts were added to 15 µL of resin and incubated at 4ºC with rotation for 2 hours. Chromotek-GFP-TRAP Beads were washed four times with IPP150 by centrifugation at 2500 x *g*. Proteins were removed by boiling in 2X LDS + 200 mM DTT or by phenol-chloroform extraction. Western blot analysis was performed as described above. Amersham ECL Plus Western blotting detection system (RPN2132) was used for quantitation by chemifluorescence, and blots were scanned using a Molecular Dynamics Storm 860 imager or GE Typhoon TRIO.

2.2.7 rRNA analysis

Whole cell RNA was prepared by lysing cells in 1:1 volume of acid phenol to AES buffer (50 mM NaOAc, pH 5.3, 10 mM EDTA, 1% SDS) at 65ºC with vortexing (Wise 1991). After chilling on ice for 5 minutes, RNA was purified using 5 Prime Phase Lock Gel tubes (5 Prime, #2302830) and as directed by manufacturer, with the addition of one organic extraction. RNA was precipitated by isopropanol and pellets washed with 70% ethanol. RNA (10 µg) was lyophilized, denatured, and separated on a 1% agarose

2.2 M formaldehyde gel as described (Sambrook and Russell 2001). RNAs were transferred to nylon membrane (Whatman) in a downward capillary transfer and northern blotted using 10 picomoles of ^{32}P-γ-labelled oligo in Church Buffer (500 mM Na_2HPO_4 pH 7.0, 7% SDS, 1 mM ETDA pH 8.0) at 42°C for 12 hours to overnight. Oligonucleotide sequences used for northern blotting were previously described (Combs *et al*. 2006) and (Table 2.2).

2.2.8 Pulse chase analysis

Cultures were grown to an OD (A_{600}) of 0.3-0.4 at 31°C in either S-MET or S-MET –URA medium supplemented with 2% sucrose. Cells were shifted to media containing 2% galactose + 1% sucrose for six hours and a low OD (A_{600}) was maintained throughout. Forty OD (A_{600}) units were harvested by centrifugation and resuspended in 12 ml of medium lacking methionine. After a 2 minute pulse with 175 μCi of [^{3}H]-L-methyl-methionine, aliquots were withdrawn at 0, 1, 2, 5, 10 and 20 min after the addition of 150 μg cold methionine. Cells were immediately frozen either with liquid nitrogen or dry ice, and RNA prepared as described above in Chapter 2.2.7. Signals were captured on a Kodak Tritium Storage screen for five days.

2.2.9 Primer extensions

Primer extension analysis was performed on U3 as described as previously described (Stevens and Abelson 2002), with the following modifications. One unit of laboratory-prepared reverse transcriptase was added per reaction with the U1 and U3 oligos described in Table 2.2, and the reactions incubated at 42°C for 1 hour. Precipitated cDNAs were resuspended in formamide loading buffer (80% formamide, 10 mM EDTA,

10 mM XC/BPB). Samples separated by electrophoresis through a denaturing gel (7% 19:1 acrylamide: bisacrylamide, 1x TBE, 8 M urea) at 25 mA for 2 hours and 30 minutes. After electrophoresis, gels were soaked in a fix solution (50% methanol 10% acetic acid) for 10 minutes, followed by 3% glycerol for 10 minutes before drying.

2.2.10 Glycerol gradient sedimentation of whole cell splicing extract

Extracts were prepared as described above; however, further clarification was achieved by centrifugation at 100,000 *x g* for 1 hour in a Beckman TLA 100.3 rotor. Protein concentration was determined by Bradford assay and 25 OD (A_{595}) units were loaded on a 10-30% glycerol gradient. Gradients were centrifuged and fractions processed for northern blotting as described in Chapter 2.2.5.

2.2.11 Native gel analysis

Extracts were prepared as described above in Chapter 2.2.9, followed by dialysis in Buffer D (20% glycerol, 20 mM HEPES, pH 7.9, 50 mM KCl, 0.5 mM DTT, 1.5 mM $MgCl_2$) twice for a total of 2 hours. Following dialysis, extracts were centrifuged at 20,000 *x g* for 10 minutes at 4°C, then frozen in liquid nitrogen and stored at -80°C. Protein concentrations were determined by Bradford assay and samples normalized to load 50 μg. Separation of native snRNPs was performed on 0.5X TAE, 4% PAGE [80:1 acrylamide:bisacrylamide] gels for at 250 volts for 4 hours at 4°C. RNAs were transferred to nylon membranes in 25 mM Na_2PO_4 and northern blotted with a oligonucleotide complementary to U6 snRNA (Combs *et al*. 2006).

2.2.12 Environmental stress conditions

All cultures were grown in YP media with 2% dextrose. For stationary phase time course, cells were harvested at OD (A_{600}) 0.3 for time 0, then at 24 and 48 hours. The heat shock time course was performed by shifting cells from 25˚C to pre-warmed media and growing at 37˚C for 20 minutes. For osmotic shock, cells were grown to OD (A_{600}) 0.5 at 31˚C then split and half shifted to YP media with 2% dextrose and 0.7 M NaCl. Both treated and non-treated cultures were returned to 31˚C for 45 minutes. All cells pellets were frozen in liquid nitrogen before proceeding with whole cell protein preparation and western blotting with chemifluorescence as described above. For a normalization control, an HRP-conjugated antibody against the endogenous H3 was utilized (Abcam, ab21054) at 1:5000 in BLOTTO.

2.3 RESULTS

2.3.1 Prp43p associates with G-patch proteins a mutually exclusive manner

Since Prp43p might function to connect in two gene expression pathways, we hypothesized that Prp43p homeostasis may be orchestrated by its interactions with three G-patch proteins, Pfa1p, Ntr1p and Gno1p. Since two of the three are known to stimulate the activity of Prp43p, this model of recruitment and activation is attractive (Lebaron *et al*. 2009; Tanaka *et al*. 2007). However, this model would require that each G-patch protein bind Prp43p in the absence of the other two so as not to create conflicting intracellular signals. Previous evidence has indicated that Pfa1p interacts with Ntr1p by the two-hybrid assay (Pandit *et al*. 2009). We sought to determine if these two, or any two cofactors, exist together in the same complex with Prp43p. We constructed a set of

strains containing a GFP tag on Prp43p and a TAP tag on one of the G-patch proteins (Fig. 2.1A; Table 2.1). In the same strain, we introduced a MYC13 tag (Longtine *et al*. 1998) into the locus of one of the other G-patch proteins, making a collection of six strains. The TAP tagged G-patch proteins were captured on IgG beads (Puig *et al*. 2001), and the presence of Prp43p or cofactors was determined by western blot. Ntr1p-TAP, Gno1p-TAP, and Pfa1p-TAP each copurified with Prp43p-GFP (Fig. 2.1B, C, and D lane 3), but did not copurify with each other (Fig. 2.1 B, C, and D lanes 6 and 9). All antibodies and tags were specific in western blotting and immunoprecipitation (Fig. 2.1E). Our data indicate that all three G-patch proteins interact with Prp43p in a mutually exclusive manner.

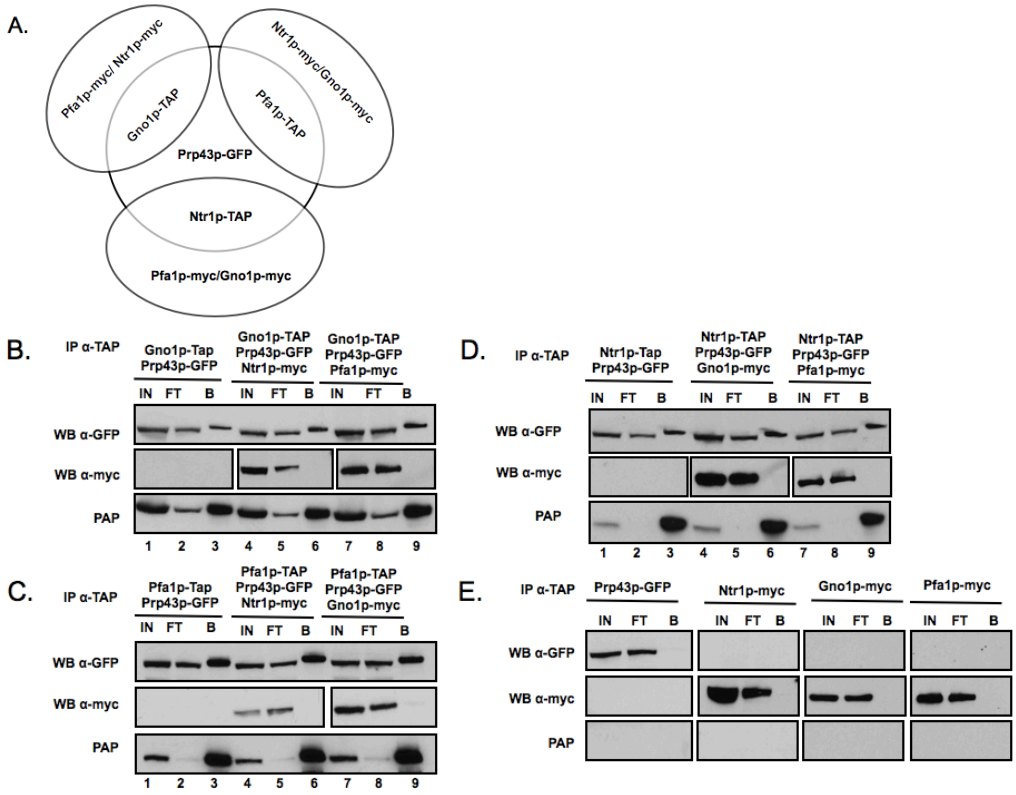


Figure 2.1 Each G-patch protein associates with Prp43p in a mutually exclusive manner.

A). Graphical representation of the experimental design. B). Purification of material associated with Gno1p-TAP. C). Purification of material associated with Pfa1p-TAP. D). Purification of material associated with Ntr1p-TAP. In the top and middle panels of B-D, coimmunoprecipitated proteins were western blotted with α-GFP (Prp43p) and α-myc (other G-patch proteins), respectively. Bottom panels demonstrate efficiency of TAP immunoprecipitation. E). Demonstration of the specificity for immunoprecipitations and antibodies used in western blotting.
IP; immunoprecipitation against the epitope tag noted; WB; Western blotting using the primary antibody named; IN, 1% input; FT, 1% unbound; B, 20% bound

2.3.2 Expression of G-patch proteins is independent of one another

Since Prp43p associates with each cofactor independently, we reasoned that they compete for Prp43p occupancy. We note that of the three, only Ntr1p is essential and the loss of Gno1p results in slow growth (Guglielmi and Werner 2002; Boon *et al*. 2006). In contrast, overexpression of Pfa1p results in cell growth arrest (Pandit *et al*. 2009). We verified these growth phenotypes by placing each protein under a *GAL1* promoter at the genomic locus of each gene (Fig. 2.2A). All proteins were depleted efficiently within 4 hours of glucose repression as determined by western blotting for the N-terminal HA-tags (Fig. 2.2B). Additionally, we utilized plasmids containing C-terminally HA/ZZ tagged *NTR1, GNO1* or *PFA1* (Open Biosystems) to overexpress each protein. Overexpression from the ectopic loci resulted in exacerbated growth defects (Fig. 2.3A). Each ectopically expressed HA/ZZ protein was produced to qualitatively similar amounts (Fig. 2.3B) and was not expressed prior to galactose induction (Fig. 2.3C). The deleterious effect of Pfa1p overexpression from either the genomic or ectopic loci is not due to the N-terminal HA or C-terminal HA/ZZ tags.

We attempted to rescue the growth defects caused by Pfa1p overexpression by simultaneously overexpressing Prp43p. If Prp43p were unable to interact with its other cofactors due to excess Pfa1p, then increasing the amount of Prp43p in the cell would bypass the defect. In a *GAL::HA-PFA1* strain carrying a p*GAL-PRP43-HA/ZZ* construct, we simultaneously overexpressed HA-Pfa1p and Prp43p-HA/ZZ. However, serial dilutions demonstrated that excess Prp43p-HA/ZZ alone results in cell growth defects, and multiple suppressor colonies were evident (Fig. 2.4A). Analyses for aberrant pre-

mRNA splicing and rRNA processing revealed no differences between wild-type and p*GAL-PRP43-HA/ZZ* (data not shown). Due to its lack of inherent substrate specificity, it is possible that overexpressed Prp43p is acting in an unregulated manner on off-pathway targets. Further investigation will be necessary to determine the molecular mechanism behind this phenotype. Whole cell RNAseq analysis may be useful in identifying which gene expression pathways are being targeted.

Work from Henry and colleagues showed that the levels of Pfa1p are reduced in a *gno1Δ* strain (Lebaron *et al*. 2009). If overexpressed Pfa1p outcompetes Gno1p and Ntr1p for Prp43p binding, then reducing Pfa1p levels may alleviate the growth defects caused by the loss Ntr1p or Gno1p. We observed that *pfa1Δ* does not alleviate the growth defects or lethality caused by the depletion of Gno1p or Ntr1p, respectively (Fig. 2.4B). This is consistent with a model where Ntr1p is necessary to bring Prp43p into the spliceosome (Tsai *et al*. 2007).

If the timing of RNA Pol I and Pol II processing is coordinated though Prp43p, cells might upregulate or downregulate these factors to influence Prp43p localization and function. To maintain homeostasis of Prp43p, the expression of the cofactors might be co-regulated. To determine if the G-patch proteins are co-regulated, we tested the expression level of one cofactor throughout time courses depleting the other two cofactors. We did not observe changes in the steady-state levels of Prp43p or its other partners during the depletion of any cofactor (Fig. 2.5). We were unable to repeat the previously published results in which a correlative relationship between Gno1p and Pfa1p was observed (Lebaron *et al*. 2009).

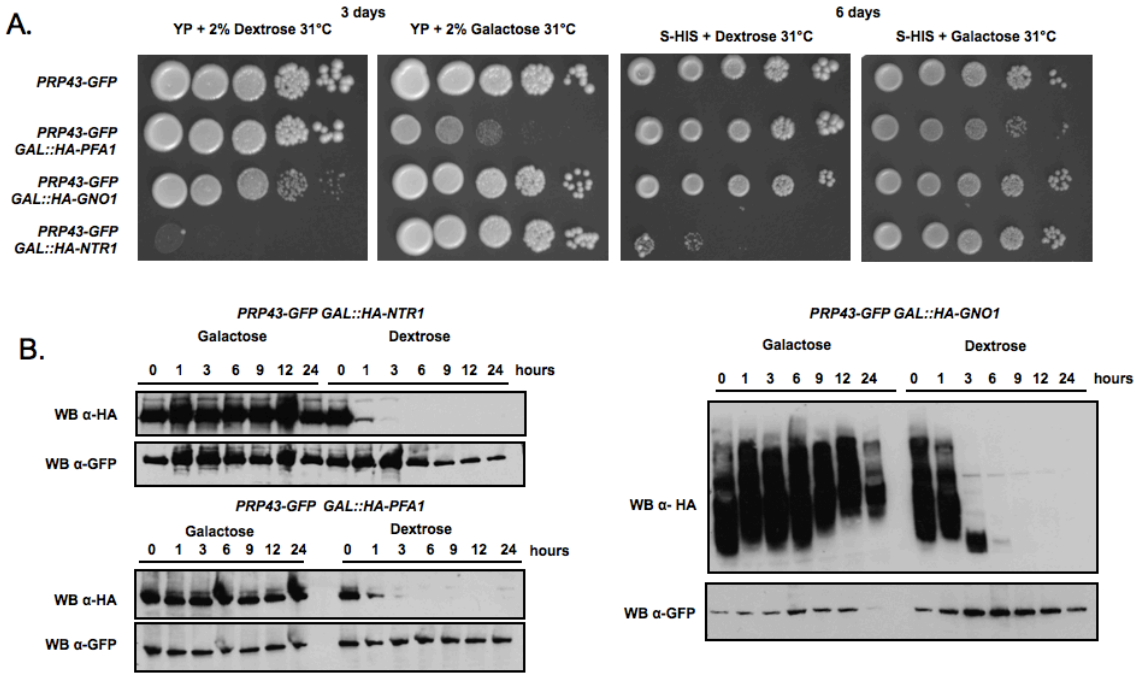


Figure 2.2 Validation of repression of G-patch proteins and their effects on cell growth.

A). Serial dilutions of genomically-encoded GAL-regulated G-patch proteins controlled plated on rich and minimal media supplemented with either dextrose or galactose. Rich media plates were photographed after 3 days at 31°C, and SD-HIS plates were photographed after 6 days at 31°C. B). Western blot analysis over a time course for genomic *GAL1-HA* tagged Ntr1p (left top panel), Pfa1p (left bottom panel) and Gno1p strains (right panel). For each time course, levels of the tagged cofactor (top) and Prp43p-GFP (bottom) are shown. Complete depletion of all cofactors is achieved within 6 hours. When analyzing total protein from yeast cells, the migration of the bands for Gno1p was consistently heterogeneous. We did not further investigate the nature of this heterogeneity.

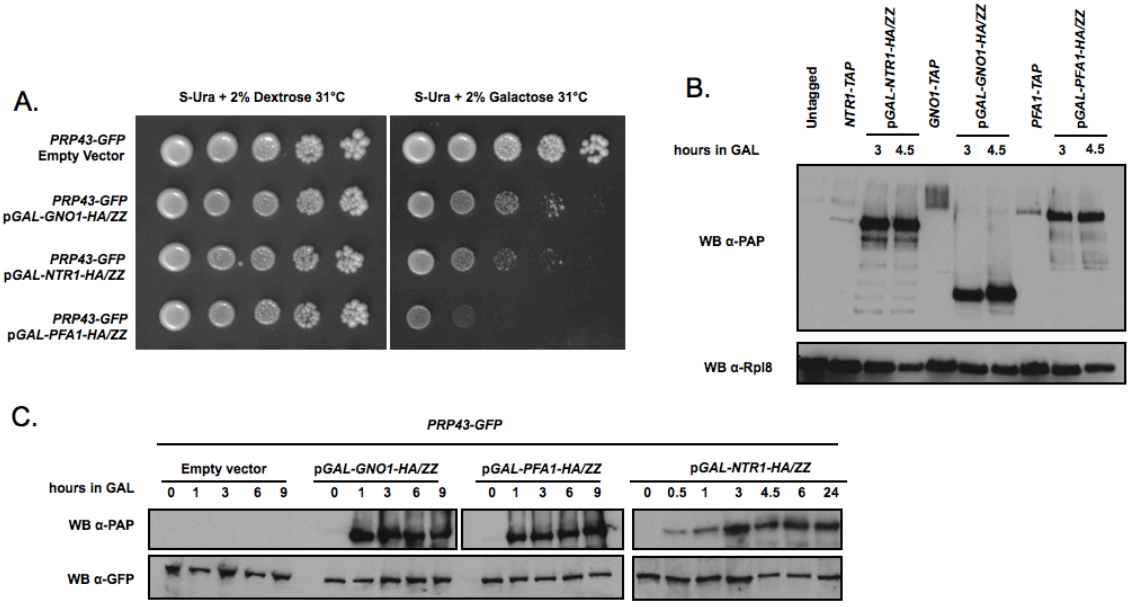


Figure 2.3 Ectopic expression of G-patch protein constructs results in growth phenotypes similar to genomic overexpressions.

A). Serial dilutions of *PRP43-GFP* strains carrying empty vector or vectors encoding the indicated C-terminally tagged HA/ZZ cofactors and plated on the indicated media at 31˚C. Plates were photographed after 3 days. B). Western blot analysis of ectopically overexpressed GAL-driven cofactors at either 3 or 4.5 hours growth in 2% galactose and 1% sucrose. For comparison, levels of the genomic TAP tagged versions of these strains for each G-patch protein are also shown. Western blotting against Rpl8p serves as a loading control. C). Western blot analysis of proteins from strains shown in (D) after shift to media supplemented with 2% galactose plus 1% sucrose demonstrates maximal overexpression within 3 hours (top panel). Western blot analysis of Prp43p-GFP is also shown (bottom panel) and indicates that total Prp43p-GFP levels do not change during the overexpression of any cofactor.

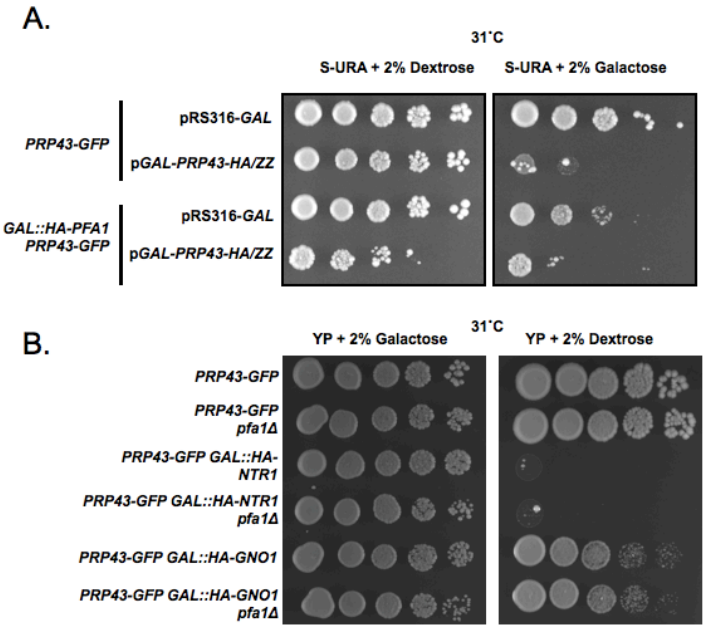


Figure 2.4 *pfa1*Δ does not alleviate the growth defects of Gno1p or Ntr1p depletion.

A). Serial dilutions of strains overexpressing Prp43p-HA/ZZ reveal growth defects. The wild-type or the genomic HA-Pfa1p depletion strain with either an empty vector or the p*GAL-PRP43-HA/ZZ* vector was grown on S-URA media with 2% dextrose or 2% galactose. Plates were incubated at 31˚C for 4 days before photographing. B). Serial dilutions of HA-GNO1 or HA-NTR1 *pfa1Δ* can not rescue either the loss of Gno1p nor Ntr1p. Strains were grown in media supplemented with galactose and serial dilutions were plated on rich media containing either 2% galactose or 2% dextrose. Plates were incubated at 31°C for 3 days before photographing.

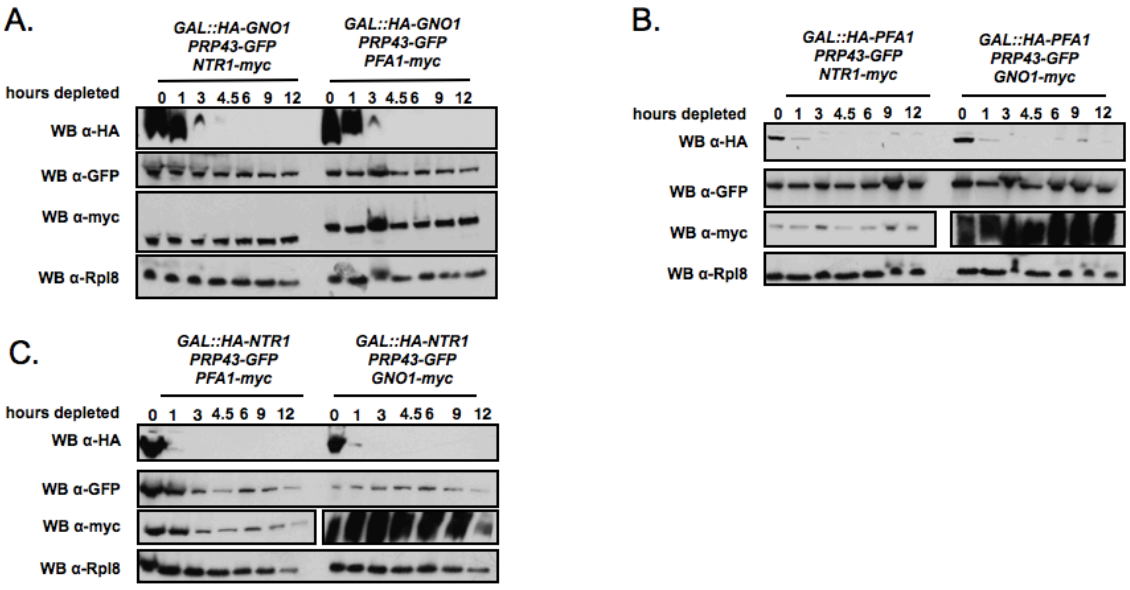


Figure 2.5 Expression of cofactors is not co-regulated.

Western blot analysis for G-patch cofactor abundance during the metabolic depletion of HA-Gno1p (A), HA-Pfa1p (B), or HA-Ntr1p (C). In each, the top α-HA western blots demonstrate the depletion of each HA-tagged cofactor. Second and third panels of each show α-GFP (Prp43p) and α-myc (other G-patch proteins) western blots respectively. An α-Rpl8p western blot is shown as a loading control. Decreases in Prp43p-GFP, Pfa1p-myc, Gno1p-myc and Rpl8p signals in the *GAL::HA-NTR1* strain are likely due to cell death that begins at approximately 6 hours after shift to non-permissive media (Boon *et al*. 2006).

2.3.3 Excess Pfa1p results in a early ribosome biogenesis defects

If excess Pfa1p inhibits the function of Prp43p in ribosome biogenesis, then pre-rRNA processing defects similar to *prp43* ATPase mutants should be evident. Previous data have shown that the loss of Pfa1p does not cause rRNA processing defects, but its overexpression decreases the levels of all rRNA precursors (Pandit *et al*. 2009; Pertschy *et al*. 2009). To identify the steps at which ribosome biogenesis is being perturbed, we analyzed the steady state rRNA levels during a time-course overexpressing Pfa1p. Northern blot analysis for the pre-rRNAs revealed that within 1 hour of inducing Pfa1p overexpression, the 35S pre-rRNA rapidly increased and persisted (Fig. 2.6A). Remarkably, the aberrant 23S pre-rRNA species also slightly accumulated, indicating that cleavages at the A_0, A_1 and A_2 sites are defective (Morrissey and Tollervey 1993). The levels of the 20S pre-rRNA decreased, consistent with the increases seen in the 35S and 23S species.

By steady state northern analysis, there was little change in the levels of the 27S pre-rRNA. The accumulation of the 35S and 23S species, and the fact that 27S pre-rRNA levels did not change, distinguish our results from those previously published (Pandit *et al*. 2009, see below). The effect is not due to the GFP tag on Prp43p as the same rRNA processing effects are seen with untagged strains harboring the p*GAL-PFA1-HA/ZZ* vector (data not shown). Depletion of Pfa1p did not result in aberrant pre-rRNA processing, similar to previous studies (Pertschy *et al*. 2009; Fig. 2.6B). We note that the pre-rRNA defects caused by Pfa1p overexpression were alleviated as Pfa1p was depleted, indicating that this Pfa1p-mediated effect is reversible.

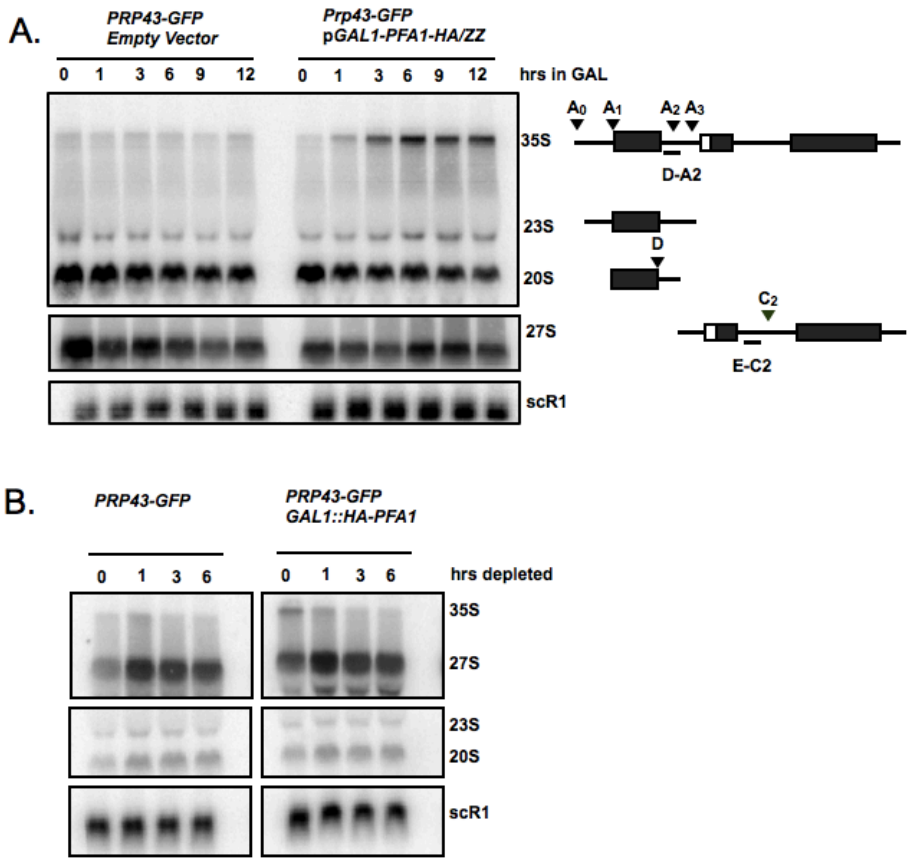


Figure 2.6 Early ribosome biogenesis defects during the overexpression of Pfa1p-HA/ZZ.

A). Steady state levels of 35S pre-rRNA and 23S rRNA increase during Pfa1p-HA/ZZ overexpression. 10 µg of total RNA were separated though a formaldehyde agarose gel, and northern blot analysis was performed using the D-A2 and E-C2 oligonucleotides (see Table 2.2) for the indicated panels. The locations of these oligonucleotides within the 35S, 23S, 20S, and 27S precursors are graphically represented and the canonical cleavage sites are indicated by triangles. sCR1 was used as an internal normalization control.
B). The aberrant rRNA processing during overexpression of Pfa1p returned to normal as Pfa1p is depleted. Northern blot analysis conducted as in (A).

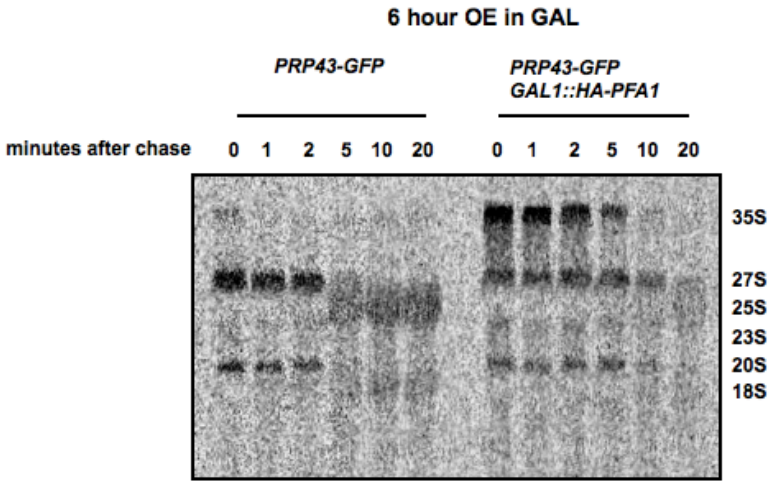


Figure 2.7 Overexpression of Pfa1p results in slowed kinetics of rRNA processing.

Pulse chase analysis of RNA from *PRP43-GFP GAL::HA-PFA1* was performed after shift to media supplemented with 1% sucrose 2% galactose for 6 hours. Cultures were pulsed with [^{3}H]-L-methyl-methionine and chased with cold methionine for the indicated times. The audioradiograph of the tritiated RNAs is shown with pre-rRNA and mature rRNAs indicated to the right.
OE, overexpressed.

We further explored effect of Pfa1p overexpression on pre-rRNA processing by pulse chase analysis with ^{3}H-L-methyl-methionine (Fig. 2.7). Under wild-type conditions, the 35S pre-rRNA was efficiently processed into 20S and 27S pre-rRNAs. These precursors were completely processed into the 25S and 18S within 5 minutes. In contrast, under conditions of Pfa1p overexpression, the accumulation of the 35S pre-rRNA observed by northern blot analysis was apparent, and the aberrant 23S pre-rRNA was also detected. Additionally, the processing of the 27S and 20S species was severely delayed with respect to wild-type, resulting in the production of very little mature rRNA within 20 minutes. Therefore, we conclude that overexpression of Pfa1p prohibits an early step in ribosome biogenesis, resulting in the slowed kinetics of rRNA processing.

2.3.4 Pfa1p overexpression abolishes Prp43p association with the 35S pre-rRNA

Mutants of Prp43p are defective in early cleavage events of ribosome biogenesis, A0, A1, and A2 (Combs *et al*. 2006). Since an early rRNA processing defect was detected by pulse chase, we hypothesized that Pfa1p may be able to inhibit the targeting of Prp43p to the 35S pre-rRNA. We performed primer extensions analysis for the 35S pre-RNA using RNA isolated from Prp43p-GFP RNA immunoprecipitations under conditions in which Pfa1p was at normal or elevated levels. Under wild-type conditions, Prp43p-GFP associated with the 35S pre-rRNA; however, if the levels of Pfa1p were increased, this association was abolished (Fig. 2.8A). Western blot for Prp43p-GFP showed similar immunoprecipitation efficiency for both conditions (Fig. 2.8B). Prp43p is not able to address the 35S pre-rRNA transcript, resulting in its accumulation and the reduction of rRNA precursors, as seen by northern blot and pulse chase analyses. We also

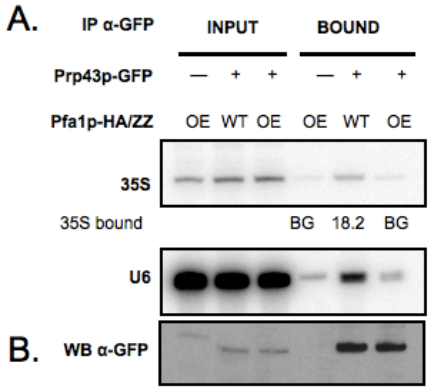


Figure 2.8. Overexpression of Pfa1p can preclude Prp43p from associating with the 35S pre-rRNA.

A). Primer extension analysis from Prp43p purified material for 35S pre-rRNA and U6 snRNA. RNA was purified by Prp43p-GFP IP from extracts containing either empty vector or p*GAL-PFA1-HA/ZZ* after shift to media supplemented with galactose for 6 hours. An untagged p*GAL-PFA1-HA/ZZ* strain (—, OE) served as a negative control for immunoprecipitation. The quantitation for the amount of 35S pre-rRNA present in the Prp43p-GFP purified material is shown below. Signals were quantitated and normalized to the negative control. 1% input; 50% bound. B). Western blot analysis of Prp43p-GFP demonstrating the efficiency of GFP immunoprecipitation. Prp43p-GFP was equally immunoprecipitated from extracts prepared under wild-type or Pfa1p-HA/ZZ overexpression conditions. 1% input; 20% bound.
OE, overexpressed levels; WT, wild-type levels

noted that Prp43-GFP did not associate with the U6 snRNA when Pfa1p was overexpressed, which indicates a concomitant loss of association with its substrates in pre-mRNA splicing (see Chapter 2.3.7).

2.3.5 Pfa1p overexpression impacts Prp43p/Gno1p association

Due to the ribosome biogenesis defects and loss of association with the 35S pre-rRNA, we investigated how Prp43p associations are affected in the presence of overexpressed Pfa1p. We hypothesized that Pfa1p affects early ribosome biogenesis by sequestering Prp43p from its potential cofactor, Gno1p. We performed coimmunoprecipitations against Prp43p-GFP and analyzed its association with Gno1p in the presence or absence of the overexpressed Pfa1p. Subsequent western blot analysis indicated that when Pfa1p was overexpressed the association between Prp43p and Gno1p was drastically reduced (Fig. 2.9, cf lanes 9 and 12). Therefore, we conclude that the increased levels of Pfa1p are able to compete with Gno1p for Prp43p association. The loss of Prp43p/Gno1p association is most likely the reason for the ribosome biogenesis defects observed. Indeed, the ribosome biogenesis defects observed during Pfa1p overexpression are similar to those reported for *gno1Δ*(Guglielmi and Werner 2002). In both instances, the 35S and 23S levels are increased and 20S levels are reduced.

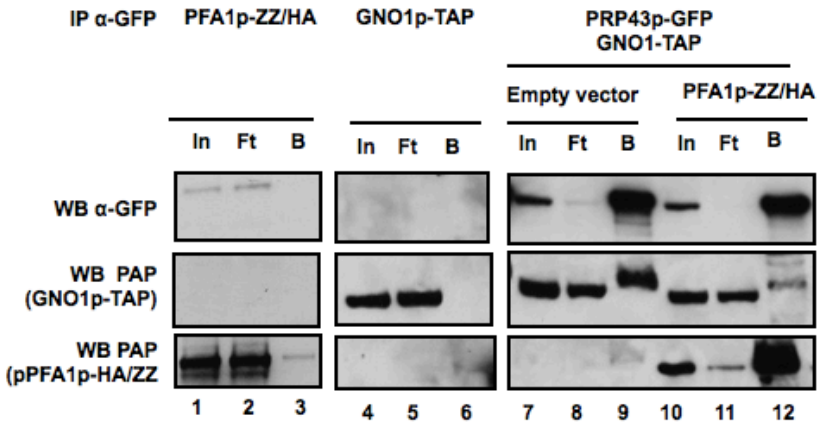


Figure 2.9 Overexpression of Pfa1p reduces Prp43p/Gno1p association.

Strains harboring either an empty vector (pRS316-*GAL*) or p*GAL1-PFA1-HA/ZZ* were shifted to media supplemented with 2% galactose and 1% sucrose for 6 hours, then Prp43p-GFP complexes were purified using α-GFP antibodies. Western blots in the top panels demonstrate Prp43p-GFP IP efficiency. Bottom panels demonstrate the overexpression of Pfa1p-HA/ZZ. Gno1p-TAP control was grown in 2% dextrose. Loading and notations are as in Figure 2.1.

2.3.6 Splicing defects are observed when Pfa1p is in excess but not when Pfa1p is absent

To further understand the molecular mechanisms underlying the growth phenotype of overexpressed Pfa1p, we investigated its relationship to Prp43p and pre-mRNA splicing. A pre-mRNA accumulation phenotype has previously been observed during the overexpression of Pfa1p (Pandit *et al*. 2009). We verified these results in our strains by primer extension analysis for the U3 pre-snoRNA during a time course of Pfa1p overexpression (Fig. 2.10A). Unpublished data from our lab suggested that Pfa1p might be a part of an uncharacterized pre-mRNA splicing subcomplex along with Prp43p. However, the loss of Pfa1p does not impact the association of Prp43p with the spliceosome nor its ability to mediate disassembly in an ATP-dependent manner (Fig. 2.10B). In ATP-depleted extracts prepared from wild-type or *pfa1Δ* backgrounds, Prp43p-TAP associated with a 16S mono-U5 snRNP and U2/U6•U5 25S-40S particles. The addition of ATP to the Prp43p-TAP *pfa1Δ* extract still resulted in the disassembly of the U2/U6•U5 particle into the individual snRNPs, U2, U6, and U5. The U2 and U6 signals peaked in fractions 10 and 4, respectively, with the addition of ATP. Therefore, we conclude that only the overexpression of Pfa1p results in aberrant splicing.

We hypothesized that the pre-mRNA splicing defect might be due to a defect in spliceosome recycling, which could result in reduced snRNAs. To further investigate the pre-mRNA splicing defect observed in the Pfa1p overexpression time courses, we analyzed whole cell snRNA distribution (Fig. 2.11A & B). Splicing extracts from strains harboring pRS316-*GAL* or p*GAL-PFA1-HA/ZZ* vectors were prepared after growth for 6

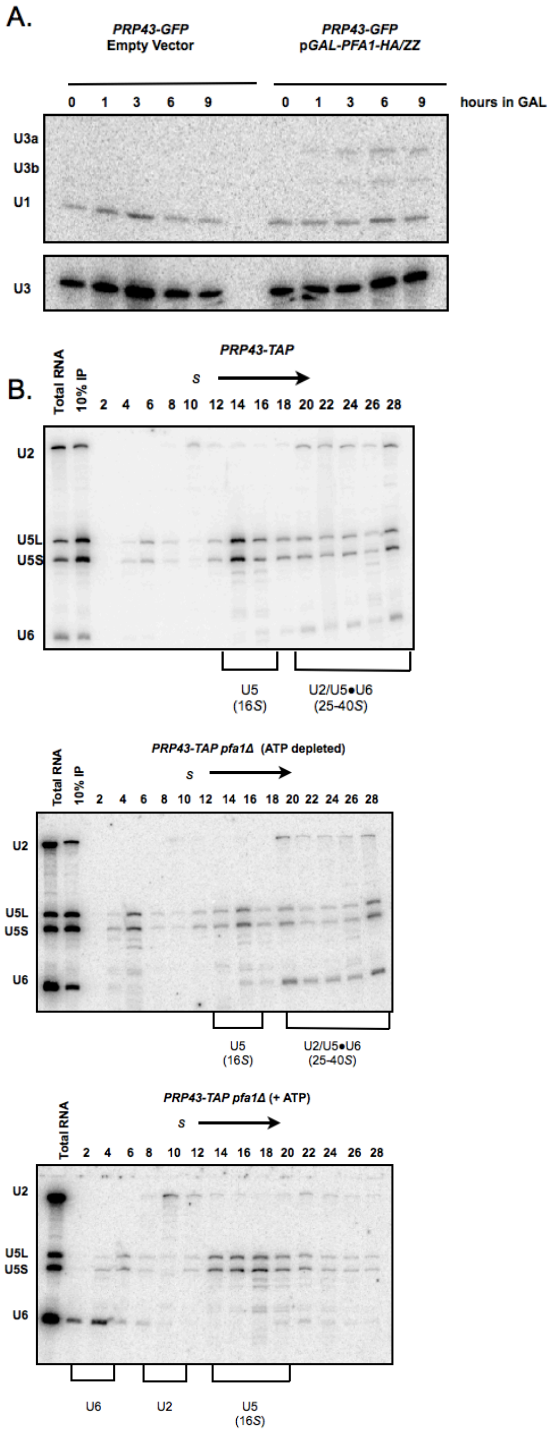


Figure 2.10 Loss of Pfa1p does not affect spliceosome disassembly.

(Figure 2.10 contd.)

A). Overexpression of Pfa1p results in a first-step pre-mRNA splicing defect. *PRP43-GFP* strains harboring an empty vector or p*GAL-PFA1-HA/ZZ* were shifted to media supplemented with 2% galactose plus 1% sucrose. RNAs were extracted at the indicated time points and subjected to primer extension analysis for pre-U3 snoRNA, U3 snoRNA or U1 snRNA. B). Prp43p-TAP associates with the spliceosome and mediates disassembly in the absence of Pfa1p. Spliceosomal particles were purified from extracts after Prp43p-TAP immunoprecipitation, then separated on a 10-30% glycerol gradient. Northern blot analysis for U2, U5 and U6 snRNAs showed Prp43p-TAP associated with both 16S U5 and U2/U6•U5 particles in either ATP-depleted wild-type (top) or *pfa1Δ* (middle) extracts. Incubating with ATP allowed the purified U2/U6•U5 snRNP to be disassociated by Prp43p-TAP even in the absence of Pfa1p (bottom).

hours in media supplemented with galactose. Separation of the whole cell extracts through 10-30% glycerol gradients revealed aberrant distribution of the U4 and U6 snRNAs in spliceosomal particles with Pfa1p-HA/ZZ overexpression. Most striking was the increase in U4 snRNA sedimenting higher in the gradient (fractions 10-12 versus 16-18) in the presence of excess Pfa1p (Fig. 2.11C, top). While the levels of U4 or U6 snRNAs were not affected, the overall levels of U5 snRNA were reduced by 44%. These characteristics have also been observed when the G-patch protein Ntr1p is depleted (Boon *et al*. 2006).

We hypothesized these shifts in U4 and U6 sedimentation reflected changes in the U4/U6 and U4/U6•U5 snRNP levels under Pfa1p overexpression conditions. We separated snRNP complexes on a native gel and visualized tri-snRNP and di-snRNP complexes by Northern blotting for the U6 snRNA. Consistent with the glycerol gradient analysis, upon overexpression of Pfa1p, we observed a stark decrease in U4/U6•U5 levels and a concomitant marked increase in U4/U6 snRNP (Fig. 2.12A). In the presence of excess Pfa1p, U4/U6 snRNP levels were 16X that of wild-type and tri-snRNP levels were decreased to 25% that of wild-type (Fig. 2.12B). We conclude that the pre-mRNA splicing defect observed during Pfa1p overexpression conditions is due to the loss of tri-snRNP. Previous results indicated that the U5 snRNP was determined to the limiting factor in tri-snRNP formation when Ntr1p was depleted (Boon *et al*. 2006). Similarly, the limiting factor in tri-snRNP formation during the overexpression of Pfa1p also appears to be the U5 snRNP.

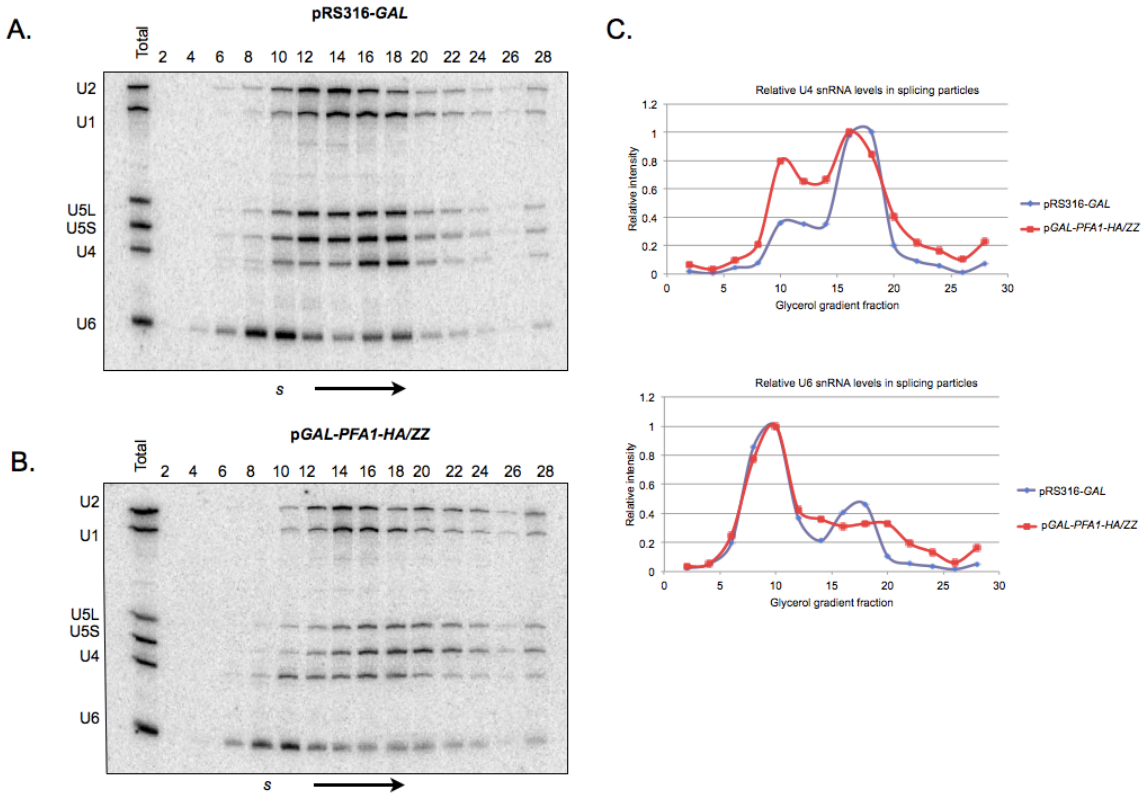


Figure 2.11 Distributions of snRNAs are disrupted in Pfa1p overexpression conditions.

A-B). Splicing extracts were prepared from strains harboring either the empty vector pRS316-*GAL* (A) or p*GAL-PFA1-HA/ZZ* (B) after growth in media supplemented with galactose for 6 hours. Whole cell splicing extracts were centrifuged though a 10-30% glycerol gradient for 16 hours at 90,500 x *g*. RNAs were extracted from alternate fractions and analyzed by Northern blotting for U1, U2, U4, U5, and U6 snRNAs, as indicated to the left. The two isoforms of U5 snRNA in *S. cerevisiae*, U5L and U5S, are noted. Equal amounts of total RNA prepared from each extract demonstrate the presence of all snRNAs in both extracts. C). Quantitation of the U4 (top) snRNA indicated an increased amount of free U4 snRNP in fraction 10 in the presence of excess Pfa1p. Additionally, a loss of U6 snRNA was observed in fractions 16 and 18 compared to the wild-type. SnRNA levels were normalized to total signal over the entire gradient.

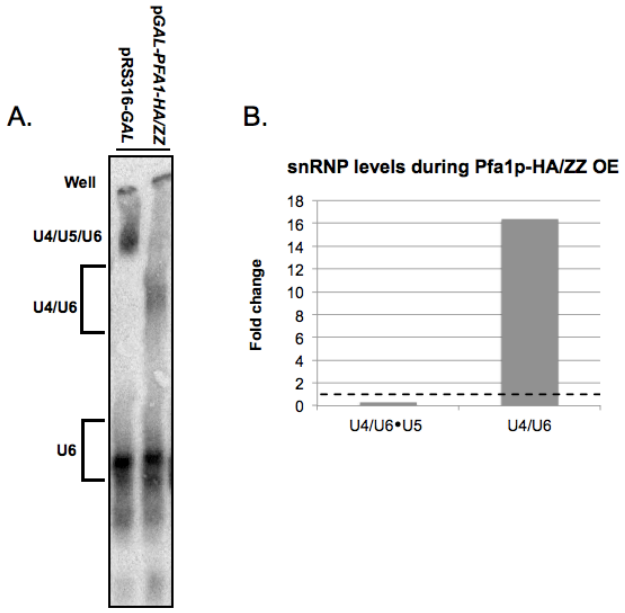


Figure 2.12 Excess Pfa1p results in decreased tri-snRNP levels.

A). Representative gel of the U4/U6 defects observed when Pfa1p is overexpressed unbalanced. Splicing extracts were prepared from strains after shift to media containing galactose for 6 hours, and 50 µg of total protein were separated on a 4% [80:1] acrylamide native gel followed by northern blot analysis for the U6 snRNA.
B). Quantitation for the increase in U4/U6 and decrease in U4/U6•U5 as seen in (A). U6 snRNA signals were normalized to the pRS316-*GAL* control. The dashed black line represents the wild-type baseline.

2.3.7 Overexpression of Pfa1p precludes Prp43p association with Ntr1p

Since the overexpression of Pfa1p resulted in pre-mRNA splicing and snRNP defects, we hypothesized that excess Pfa1p affects the association between Prp43p and its spliceosomal cofactor, Ntr1p, in a manner similar to Gno1p (Fig. 2.9). We analyzed the association of Prp43p-GFP with Ntr1p in the presence of overexpressed Pfa1p by coimmunoprecipitation. Subsequent Western blot analysis indicated that when Pfa1p was overexpressed the association between Prp43p and Ntr1p was drastically reduced (Fig. 2.13, cf. lanes 15 and 21). Therefore, we conclude that the increased levels of Pfa1p are able to outcompete Ntr1p for association with Prp43p, in a manner similar to Gno1p. The loss of Prp43p/Ntr1p association is most likely the reason for the snRNP assembly defects and downstream pre-mRNA accumulation. Therefore, Pfa1p has the ability to sequester Prp43p away from the other two G-patch proteins with which it associates.

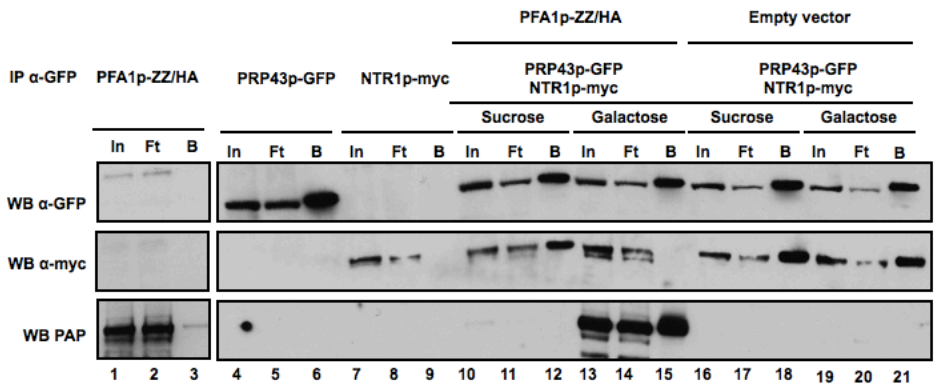


Figure 2.13 Prp43p loses association with Ntr1p when Pfa1p is in excess.

Coimmunoprecipitations followed by serial Western blotting for the indicated tags was completed using extracts from strains either an empty vector (pRS316-*GAL*) or p*GAL1-PFA1-HA/ZZ* shifted to media supplemented with 2% galactose and 1% sucrose for 6 hours. Western blots in the top panels demonstrate Prp43p-GFP IP efficiency. Bottom panels demonstrate the overexpression of Pfa1p-HA/ZZ. Control strains Prp43p-GFP and Ntr1p-myc were grown in 2% dextrose, while the p*GAL-PFA1-HA/ZZ* only control was grown in 2% galactose 1% sucrose. Loading and notations as in Figure 2.1.

2.3.8 *PFA1* is predicted to connect to ribosome biogenesis or pre-mRNA splicing solely via *PRP43*

We hypothesized that if Pfa1p has the ability to preclude Prp43p from pre-mRNA and pre-rRNA processing pathways, it is likely to be regulated differently than either Gno1p or Ntr1p. Perhaps in response to environmental stresses, Pfa1p levels are adjusted to halt or pause both of the RNA processing pathways in which Prp43p is involved. Using the YeastNet 2.0 database (Lee *et al*. 2007), we calculated the probability of *bona fide* linkages between genes related to the G-patch proteins. YeastNet 2.0 uses parameters such as, but not limited to, co-expression microarray data, mass spectrometry data, and yeast two-hybrid interactions to calculate an overall score for each predicted interaction.

The log-likelihood scores for the interactions of *GNO1*, *PFA1*, and *NTR1* were used to generate a predicted network prepared in Cytoscape (Shannon *et al*. 2003). Our network indicates that *GNO1* and *NTR1* lie in rRNA processing and pre-mRNA splicing clusters, respectively (Fig. 2.14). Node size is dependent upon the score, with more significant scores being larger. The cluster surrounding *GNO1* contains both large and small ribosomal subunit biogenesis factors, including many 90S processing factors, such as Mpp10 and multiple UTPs (Dragon *et al*. 2002). The pre-mRNA splicing and ribosome biogenesis clusters are connected by two nodes, *PRP43* and *CWC22*. *CWC22* is an essential pre-mRNA splicing factor, and its presence is required for the ATP-dependent function of Prp2p (Yeh *et al*. 2011) and the removal of the SF3a/b complexes prior to the first step of splicing (Lardelli *et al*. 2010). The edge linking *CWC22* to *GNO1* is based on the relatively poorly correlated co-citation score, while the link to *NTR1* was

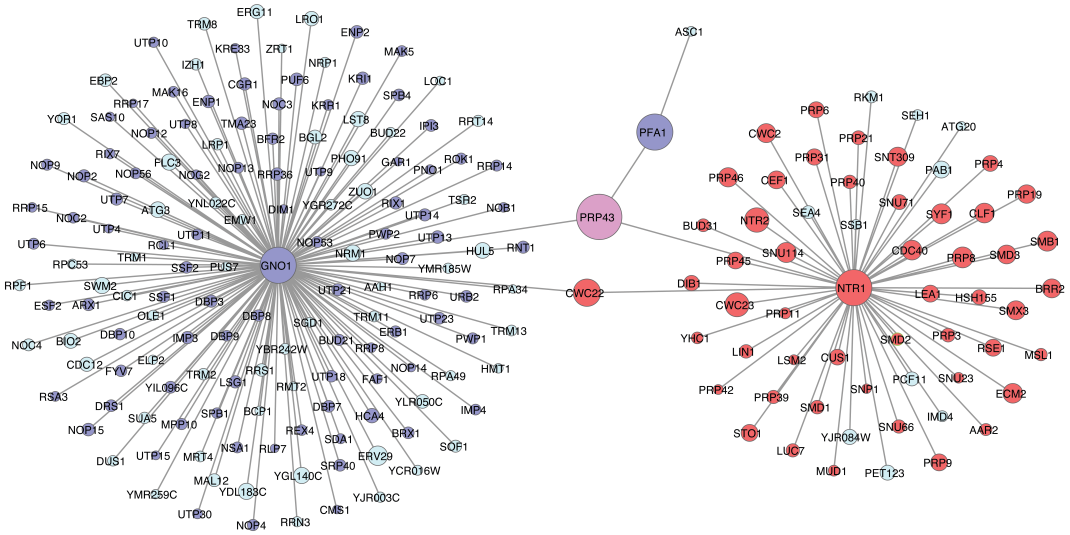


Figure 2.14 The regulatory network connections predicted for *PFA1* lie outside the pre-mRNA splicing and ribosome biogenesis clusters.

A predicted network of *GNO1*, *PFA1*, and *NTR1* was assembled from the YeastNet 2.0 database. *NTR1* and *GNO1* lie in pre-mRNA or pre-rRNA processing clusters, respectively, while *PFA1* is connected to these processes only through *PRP43*. Pre-mRNA processing (red nodes) and rRNA processing (blue nodes) are linked to each other by *PRP43* (purple), is involved in both processes. Light blue nodes are either uncharacterized genes or genes not known to contribute to either pathway.

determined by mass spectrometry. Strikingly, *PFA1* is connected only to ribosome biogenesis and pre-mRNA splicing via *PRP43*, indicating that it is likely to be regulated differently than other ribosome biogenesis factors.

2.3.9 Whole cell Pfa1p levels do not increase in response to environmental stress despite increases in mRNA levels

Systematic microarray experiments suggest that *PFA1* transcripts, but not *GNO1* or *NTR1* transcripts are significantly increased (≥50%) in response to hyper-osmotic shock with NaCl treatment (Berry and Gasch 2008). *PFA1* transcript levels also rise during heat shock at 37°C and during the first two days of stationary phase growth (Gasch *et al*. 2000). Other data sets indicate that *PFA1* transcript levels are increased in response to ammonium, phosphate, or amino acid starvation (Gasch *et al*. 2000; Brauer *et al*. 2008). Together, these data suggest that Pfa1p levels increase in response to environmental stress in a differential manner from Ntr1p or Gno1p. However, despite these transcriptional differences, we found that whole cell protein levels of Pfa1p did not significantly change during hyper-osmotic shock, stationary phase growth or heat shock (Figure 2.15) We noted that under heat shock a consistent shift in the molecular weight of Pfa1-myc was evident, though it has yet to be determined whether this represents a *bona fide* post-translational modification or if such a modification is biologically relevant. It is also possible that the presence of increased levels of Pfa1p mRNA serves as a buffer to ensure that levels of Pfa1p remain high to enforce Prp43p association. Therefore, how the transcriptional changes described above manifest in functional differences is still unknown.

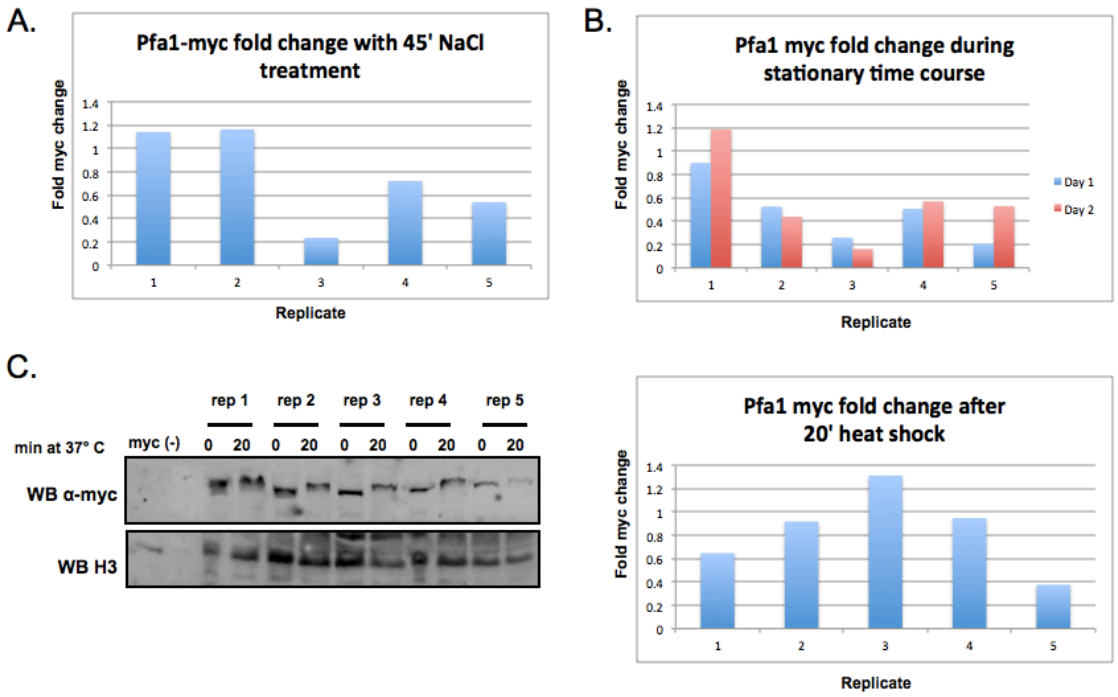


Figure 2.15 Pfa1p-myc levels do not change in response to stress.

A). Whole-cell Pfa1p-myc levels do not increase following hyper-osmotic shock with 0.7 M NaCl for 45 minutes. MYC signals from five biological duplicates were visualized and quantitated by western blotting followed by chemifluorescence. All signals were normalized to an H3 loading control, and fold change from time 0 was calculated. B). Whole-cell Pfa1p-myc levels do not increase during stationary phase. Time courses beginning at OD (A_{600}) 0.3 were conducted over 2 days allowing cultures to reach saturation. Time points at 24 and 48 hours were analyzed for total Pfa1p-myc levels as in (A). C). Cells undergoing heat shock do not upregulate the level of Pfa1p-myc. After shifting from 25˚C to 37˚C for 20 minutes, whole cell protein was isolated from five biological replicates (rep 1-5), and Pfa1p-myc levels were visualized and quantitated by western blotting followed by chemifluorescence. Though Pfa1p-myc levels did not increase, we did observe a consistent shift in molecular weight at 37˚C, perhaps suggesting some large post-translational during heat shock. Quantitation of Pfa1-myc signals by chemifluorescence is graphed to the right. Signals normalized as in (B).

2.4 DISCUSSION

Due to the multifunctional nature of Prp43p and the spatial separation between the two processes in which it functions, we set out to understand how Prp43p maintains proper distribution between pre-mRNA splicing and ribosome biogenesis, as well as how Prp43p may help to coordinately regulate both of these processes. Our work supports a working model for Prp43p in which its substrate specificity, activation, and cellular distribution is coordinated through the efforts of the three G-patch proteins, Gno1p, Ntr1p, and Pfa1p. We demonstrate that the G-patch proteins interact with Prp43p in mutually exclusive complexes, which would enforce substrate specificity and/or correct spatiotemporal activation of Prp43p in the proper pathway (Fig. 2.1). Additionally, despite previously published yeast two-hybrid data for Pfa1p and Ntr1p (Pandit *et al*. 2009), we find that none of the cofactors stably associate with one another.

Others have suggested that a competition exists between Pfa1p and the other G-patch proteins for Prp43p interaction (Pandit *et al*. 2009). Here we show that excess Pfa1p inhibits both pre-mRNA splicing and ribosome biogenesis pathways directly by outcompeting both Gno1p and Ntr1p for Prp43p binding (Figs. 2.9 & 2.13, respectively). The depletion of Ntr1p affects the regeneration of snRNPs, particularly the U5 snRNP, so that a reduction in the U4/U6•U5 tri-snRNPs and a subsequent accumulation of pre-mRNA are observed (Boon *et al*. 2006; see Chapter 3.2). We relate these phenotypes to the interaction of Prp43p with Ntr1p. Under Pfa1p overexpression conditions in which Prp43p does not interact with Ntr1p, a similar snRNP recycling defect is present (Figs. 2.11 & 2.12), consistent with the resulting pre-mRNA accumulation (Pandit *et al*. 2009,

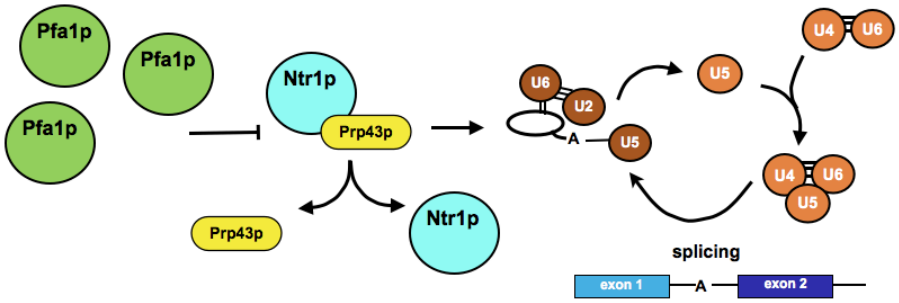


Illustration 2.1 pre-mRNA splicing defects are an indirect effect of excess Pfa1p.

Under normal conditions Ntr1p association with Prp43p allows for the dissociation of the post-splicing lariat RNP. This allows for the recycling of the snRNPs so that the U5 snRNP rejoins the U4/U6 di-snRNP to make U4/U6•U5 tri-snRNP. In the case of Ntr1p depletion, the cellular levels of U5 snRNP decrease (Boon *et al*.2006), resulting in less U4/U6•U5 since the U5 snRNP is the limiting reagent. Similarly, if cellular Pfa1p levels increase, the Ntr1p/Prp43p interaction is lost so that U5 is not properly recycled, indirectly resulting in a halt in pre-mRNA splicing.

Fig. 2.10, Illustration 2.1). The ribosome biogenesis defects observed under Pfa1p overexpression conditions are unlikely to solely be indirect effects of defective pre-mRNA splicing, because Prp43p/35S and Prp43p/Gno1p pre-rRNA associations are abrogated (Figs. 2.8 & 2.9). We show that Pfa1p has the ability to titrate Prp43p away from its other two cofactors and, thus, their respective RNA processing pathways. Therefore, maintaining the proper balance of G-patch proteins is essential for the proper cellular distribution and function of Prp43p.

In situations where cellular stresses are encountered, the modulation of Pfa1p abundance seen by microarray experiments (Gasch *et al*. 2000; Brauer *et al*. 2008; Berry and Gasch 2008) may be the means by which cells are able to disconnect Prp43p from both ribosome biogenesis and pre-mRNA splicing, effectively reversibly stalling both processes. However, we were unable to relate the reported increases in *PFA1* mRNAs to the whole cell protein levels of Pfa1p during hyper-osmotic shock, stationary phase growth or heat shock (Fig. 2.15). Perhaps some other level of regulatory mechanism contributes to the ability of Pfa1p to pause gene expression, such as its post-translational modification (Fig. 2.15C). Therefore, it is not yet known how the transcriptional changes previously reported manifest in functional differences in response to environmental stresses.

The unique role for Pfa1p in regulating Prp43p is supported by the fact that its binding interface with Prp43p is not dependent on its G-patch domain (Pandit *et al*. 2009; Lebaron *et al*. 2009) unlike that of Ntr1p (Tanaka *et al*. 2007) or Spp2p in the case of Prp2p (Silverman *et al*. 2004). While other examples of multiple cofactors regulating a

single DExH/D-box helicase are not known, two DEAD-box helicases, Dbp5p and Ded1p are regulated by a single factor, Gle1p. Like the Prp43p/Ntr1p and Pfa1p interfaces, there is some commonality to the Gle1/Dbp5 or Ded1p binding surfaces, but genetic data suggest partially distinct regions on Gle1p as well (Bolger and Wente 2011). It is yet to be determined whether Ntr1p or Gno1p has exclusive binding surfaces on Prp43p; however, considering the extensive contacts that Pfa1p makes with Prp43p, surfaces exclusively binding only one G-patch protein may be unlikely (Walbott *et al*. 2010). Also, if each G-patch protein can modulate Prp43p activity, then it might be expected for all of them to interact with the helicase core or putative regulatory 5' hairpin in the RecA2 domain of Prp43p (He *et al*. 2010). In contrast to Ntr1p and Pfa1p, which both can activate Prp43p, Gle1p functions as an activator with Dbp5p in mRNA export and an inhibitor of Ded1p dependent start site scanning of mRNAs for translation initiation. Its specificity as an activator or inhibitor is thought to be determined by unique interactions with factors that function at the particular step of gene expression. For example, Gle1p docks to specific sites in the nuclear pore (e.g. Nup42p) during Dbp5p-mediated mRNA export and with eIF3 during Ded1p-mediated translation initiation (Bolger and Wente 2011). Both cases demonstrate opposite aspects of how helicase activity can be spatially regulated for a particular step in gene expression. Also, while cofactors of Prp43p contain G-patch domains, Gle1p does not.

As more factors that impact the activities of DExH/D-box helicases are characterized, we anticipate further examples of helicase regulation will be uncovered, particularly in higher metazoans. Already, TFIP11 has been demonstrated to be the

functional homolog of Ntr1p (Yoshimoto *et al*. 2009), and the G-patch containing factor GPKOW influences hPrp2p and hPrp16p activity (Hegele *et al*. 2012). It will be interesting to determine if other G-patch containing proteins in metazoans are also involved in regulating the activity of a helicase by stimulation and/or recruitment. Therefore, our understanding of how G-patch proteins govern the substrate specificity, activation, and cellular distribution of Prp43p in yeast will shed light on mechanisms of general DExH/D helicase function and regulation.

Chapter 3: Ntr1p and Gno1p regulation of Prp43p

3.1 BACKGROUND

Prp43p interacts with three cofactors, Pfa1p, Gno1p and Ntr1p, all of which contain a G-patch domain (Lebaron *et al*. 2005). Pfa1p is nonessential, and loss of Pfa1p does not impact cell growth under normal conditions (Pertschy *et al*. 2009). In contrast, Ntr1p is essential, and cell growth halts after 9 hours of Ntr1p depletion (Fig. 2.2; data not shown). Also, the loss of Gno1p results in severely slowed growth (Fig. 2.2; Guglielmi and Werner 2002). Unlike Pfa1p, the G-patch domains of Ntr1p and Gno1p are both required for proper cell growth (Pandit *et al*. 2009). Single point mutations in the G-patch domains of either Ntr1p or Gno1p lead to growth defects that phenocopy strains in which they are depleted or deleted, respectively (Tanaka *et al*. 2007; Guglielmi and Werner 2002). We further investigated the balance of Prp43p between ribosome biogenesis and pre-mRNA splicing in the context of Ntr1p and Gno1p to construct a complete network for Prp43p regulation.

3.1.1 Ntr1p

Ntr1p (Nineteen complex Related) was first determined to be loosely associated with Prp19p, a component of the NTC (Tsai *et al*. 2005; see 1. 2.4). Ntr1p forms a dimer with its partner Ntr2p, and both proteins are necessary for the release of the lariat intron after pre-mRNA splicing. However, neither plays a direct a role in the pre-mRNA splicing reaction. These two factors join with Prp43p to form the trimeric NTR complex. However, while Ntr1p and Ntr2p are always found as partners, the majority of Prp43p is not found as a part of this complex. Ntr1p mediates the formation of the NTR complex by

interacting with Prp43p via its G-patch domain and interacting with Ntr2p via a central domain. No direct interaction between Prp43p and Ntr2p is evident (Tsai *et al*. 2005). The NTR complex associates with the spliceosome in an ATP-independent manner via an Ntr2p-dependent interaction with the U5 snRNP (Tsai *et al*. 2007). Once associated, the NTR complex disassembles the post-splicing lariat RNP and releases the individual U2, U5, and U6 snRNPs (Tsai *et al*. 2005; Fourmann *et al*. 2013).

A single mutation in the G-patch domain of Ntr1p (L68A, L69A or L80A) is sufficient for its loss of association with Prp43p, and, therefore, the helicase activity of Prp43p can not be stimulated. Surprisingly, *ntr1L68A* is not deficient for growth at 30°C, but is temperature sensitive. However, *ntr1L68A* is synthetically lethal in combination with *prp43Y402A*. Both K398A and Y402A mutations, which lie in the β-hairpin of Prp43p, cause temperature sensitivity in Prp43p (Tanaka *et al*. 2007; He *et al*. 2010; Walbott *et al*. 2010). The prp43Y402Ap mutant loses interaction with wild-type Ntr1p *in vitro*, demonstrating that a direct interaction surface lies between Prp43p and Ntr1p lies near the β-hairpin. Moreover, in the presence of wild-type Ntr1p, the Y402A mutant is only partially able to release the lariat intron from post-spliceosomal complexes. The L68A mutation in Ntr1p renders it completely unable to work with the Y402A Prp43p mutant to disassociate the lariat intron (Tanaka *et al*. 2007).

A truncation in the CTD of Prp43p (aa 722-767) is also synthetically lethal with *ntr1L68A*, and alanine mutations in the three amino acids ^{717}WLI^{719} are sufficient for this phenotype. Simultaneous mutation of Prp43p at Y402 and deletion of the C-terminal 35 amino acids also results in synthetic sickness. Together these data suggest that Ntr1p

redundantly interacts with Prp43p at Y402A in the tip of the β-hairpin and also within its CTD (Tanaka *et al*. 2007; He *et al*. 2010; Walbott *et al*. 2010).

A mutation in the conserved threonine of motif I in Prp43p (T123A) results in a dominant negative mutant that retains the ability to bind the spliceosome but is ATPase deficient and unable to unwind the intron lariat from post-splicing lariat RNP (Martin *et al*. 2002; Tanaka and Schwer 2006). Paradoxically, if residues ^{684}IRK^{650} or ^{717}WLI^{719} in the CTD of Prp43p are replaced with alanines in the T123A mutant, the dominant negative phenotype can be relieved (Tanaka *et al*. 2007). Considering the OB-fold of the CTD is important for RNA binding (Walbott *et al*. 2010), the dominant lethality caused by the T123A mutation is most likely relieved because the mutant protein does not bind its RNA substrates. Additionally, since Prp43p also serves a purpose in ribosome biogenesis (Combs *et al*. 2006; Lebaron *et al*. 2005; Leeds *et al*. 2006), these mutations would be expected to relieve the effect of the T123A mutation in that pathway as well.

3.1.2 Gno1p

Gno1p is one of the three G-patch proteins that was shown to copurify with Prp43p (Lebaron *et al*. 2005). The G-patch domain of Gno1p is contained between amino acids 25-70 of its N-terminus, and contains 6 glycines that are conserved with its human homolog, PinX1. Additionally, Gno1p contains KK(E/D) repeats in its C-terminus, which are common in other rRNA processing factors. A deletion of Gno1p renders cells slow growing, and loss of the G-patch domain is sufficient for this slow growth phenotype. Moreover, a single point mutant in a conserved tryptophan of the G-patch domain (W38S) produces a slow growth phenotype comparable to the null mutant. It does not

appear that the KK(D/E) repeats play as an important role since truncation of Gno1p results in a milder growth defect (Guglielmi and Werner 2002).

The first indication that Gno1p plays a role in ribosome biogenesis was its localization to the nucleolus. When cells are deficient in Gno1p, the steady state levels of the 35S pre-rRNA increase and an aberrant 23S rRNA species accumulates. These results are indicative of impaired A0, A1, and A2 cleavage events (Morrissey and Tollervey 1993), which are the same cleavage events that *prp43* ATPase mutants affect (Combs *et al*. 2006; Lebaron *et al*. 2005; Leeds *et al*. 2006). Specifically, the 27S A2/A3 cleavage ratio is decreased indicating that the loss of Gno1p affects the A2 cleavage site and therefore the upstream sites are affected. The KK(D/E) repeats are not needed for the rRNA maturation step, but were shown to be important for the proper trimming of the 3' end of the U18 snoRNA. In contrast, the G-patch domain was shown to be necessary for both the trimming of U18 and rRNA maturation (Guglielmi and Werner 2002).

Though the human homolog of Gno1p, PinX1p, has been shown to play a role in repressing telomerase activity and maintaining telomere length (Zhou and Lu 2001; see 5.2), Gno1p does not appear to have this function in yeast. The sequence responsible for the telomerase function in humans maps to the C-terminus of PinX1p, which is divergent from the yeast homolog. In contrast to humans, *gno1Δ* results in a mild shortening of telomeres. The G-patch domain appears to be responsible for this phenotype since a single mutation, W38S, cannot rescue the phenotype in a *gno1Δ* strain. It is speculated that the shortened telomeres are actually an indirect effect of slow growth caused by *gno1Δ* (Guglielmi and Werner 2002). The phenotype has also been observed in DAmP

(Decreased Abundance by mRNA Perturbation) alleles of many other nucleolar genes, such as *POP1* a component of RNase MRP, exhibit shortened telomeres. Surprisingly, telomeres are also shortened in the DAmP alleles of *PRP22*, *PRP43*, and *NTR2* (Ungar *et al*. 2009). However, a mechanistic understanding of the connections between spliceosome disassembly or ribosome biogenesis with telomere maintenance in yeast has yet to be elucidated.

Gno1p interacts with Prp43p by yeast two hybrid (Lebaron *et al*. 2005); and unpublished results indicate that Gno1p can directly interact with Prp43p *in vitro* (unpublished, Y. Henry, personal communication). Gno1p sediments with 90S pre-ribosomal particles in sucrose gradients, though it has not been shown to directly associate with the 35S rRNA. Its localization to the nucleolus and the effects on A0, A1 and A2 cleavage in *gno1Δ* implicate its presence there. The rRNA processing defects are not due to a reduction in the steady state levels of Prp43p in the absence of Gno1p (Lebaron *et al*. 2009). However, interestingly, Gno1p appears to be connected to Pfa1p. A previous report showed that Pfa1p is truncated and its levels are depleted in a *gno1Δ* strain. In contrast, Gno1p levels are not affected in a *pfa1Δ* strain. The steady state levels of other 40S pre-ribosomal components are not affected, so this effect is specific to Pfa1p and not due to downstream destabilization of pre-40S particles (Lebaron *et al*. 2009). Our attempt to investigate the co-regulation of the three G-patch protein factors, however, indicated that there is no change in their steady state levels in response to the depletion of another cofactor (Fig. 2.5). Additionally, it was reported that there is an increase in the steady state levels of the 40S component Ltv1p in a *gno1Δ* strain. The upregulated Ltv1p

resulted in an increased in Ltv1p/20S pre-rRNA association. Moreover, the truncated version of Pfa1-TAP in this *gno1Δ* strain does not interact with Prp43p or the 20S pre-rRNA. It is possible that in the *gno1Δ* strain, the truncated Pfa1p, which does not associate with 20S or Prp43p, results in the stalling of Ltv1p on pre-40S particles (Lebaron *et al*. 2009). This interpretation is consistent with the genetic interaction between *LTV1* and *PFA1* (Pertschy *et al*. 2009; see Chapter 2.1.1).

3.2 MATERIALS AND METHODS

3.2.1 Yeast strains and plasmids

A full listing of yeast strains used in Chapter 3 are listed in Table 3.1. *GNO1-13MYC-TRP1* was amplified from strain SS3160 (Table 4.2) from genomic DNA and cloned into pRS316. Mutagenesis of *GNO1* was performed using oligonucleotides containing the W38S mutation (Ho *et al*. 1989; Quikchange, Agilent Technologies), and the mutation was verified by DNA sequencing. Oligonucleotides for Gno1p cloning and tagging of proteins are listed in Table 3.2. Plasmid transformation of yeast was performed as previously described (Gietz and Woods 2002).

3.2.2 Primer extension analysis of RNA

Primer extension analysis was performed on U3 as described as previously described (Stevens and Abelson 2002), with the modifications outlined in Chapter 2.2.8.

Table 3.1 Yeast strains used in Chapter 3

Strain	Relevant Gene(s)	Genotype	Source/Reference
SRY 328	*prp2-1* isogenic wild-type strain	*MAT α ade2-101, his3-d300, lys2-801, ura3-52*	(Ruby *et al*. 1993)
599	*PRP43-GFP*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX*	(Combs *et al*. 2006)
3043	*prp2-1*	*MAT α prp2-1, ade2-1, his3-352, trp1-289 ura3-52*	(Ruby *et al*. 1993)
3072	*GNO1-TAP*	*MAT **a** his3 Δ 1, leu2 Δ 0, met15 Δ 0, ura3 Δ 0, GNO1-TAP::HIS3*	(Howson *et al*. 2005)
3099	*PFA1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, PFA1-13myc::URA3*	This study
3103	*NTR1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, NTR1-13myc::URA3*	This study
3107	*GNO1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, GNO1-13myc::URA3*	This study
3129	pRS316-*GAL*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX, pRS316-GAL*	This study
3130	p*GAL-GNO1-HA/ZZ PRP43-GFP*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX*, p*GAL-GNO1-HA/ZZ-URA*	This study
3132	p*GAL-NTR1-HA/ZZ PRP43-GFP*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX*, p*GAL-NTR1-HA/ZZ-URA*	This study
3139	pRS316-*GAL PRP43-GFP NTR1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX, NTR1-13myc::TRP1, pRS316-GAL*	This study
3140	p*GAL-GNO1-HA/ZZ PRP43-GFP NTR1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX, NTR1-13myc::TRP1*, p*GAL-GNO1-HA/ZZ-URA*	This study

Strain	Relevant Gene(s)	Genotype	Source/Reference
3141	p*GAL-GNO1-HA/ZZ* *PRP43-GFP* *PFA1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX, PFA1-13myc::TRP1*, p*GAL-GNO1-HA/ZZ*	This study
3142	pRS316-*GAL* *PRP43-GFP* *PFA1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX, PFA1-13myc::TRP1*, pRS316-*GAL*	This study
3143	p*GAL-NTR1-HA/ZZ* *PRP43-GFP* *PFA1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX, PFA1-13myc::TRP1*, p*GAL-NTR1-HA/ZZ*	This study
3144	*GAL::HA-NTR1* *PRP43-GFP*	*MATα his3Δ1 ,leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, HIS3:: GAL::HA-NTR1, PRP43-GFP::KANMX*	This study
3145	*GAL::HA-GNO1* *PRP43-GFP*	*MATα his3Δ1 ,leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, HIS3::GAL::HA-GNO1, PRP43-GFP::KANMX*	This study
3157	pRS316-*GAL* *PRP43-GFP* *GNO1-TAP*	*MAT **a** his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, GNO1-TAP::HIS3, PRP43-GFP::KANMX*, pRS316-*GAL*	This study
3158	p*GAL-NTR1-HA/ZZPRP43-GFP* *GNO1-TAP*	*MAT **a** his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, GNO1-TAP::HIS3, PRP43-GFP::KANMX*, p*GAL-NTR1-HA/ZZ*	This study
3169	pRS411 *PRP43-GFP*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX*, pRS411	This study
3180	pRS411 *PRP43-GFP* *GAL::HA-NTR1*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX, HIS3:: GAL::HA-NTR1*, pRS411	This study
3211	*GAL::HA-GNO1* *PRP43-GFP* pRS416-*GNO1-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX, HIS3::GAL::HA-GNO1*, pRS416-*GNO1-myc*	This study

Strain	Relevant Gene(s)	Genotype	Source/Reference
3212	*GAL::HA-GNO1* *PRP43-GFP* pRS416-*gno1 W38S-myc*	*MATα his3Δ1, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0, PRP43-GFP::KANMX, HIS3::GAL::HA-GNO1*, pRS416-*gno1W38S-myc*	This study
3213	p*GAL-GNO1-HA/ZZ*	*MAT **a** his3Δ1, leu2Δ0, met15Δ0, ura3Δ0*, p*GAL-GNO1-HA/ZZ*	This study
3214	p*GAL-NTR1-HA/ZZ*	*MAT **a** his3Δ1, leu2Δ0, met15Δ0, ura3Δ0*, p*GAL-NTR1-HA/ZZ*	This study

Table 3.2 Oligonucleotides used in Chapter 3

Name	Sequence
Ntr1 R1	5'TCTAATCAATTTTGTTTTTCGACAATAATATATAAATCGTGCCTATCTCATCTG GAGGAAGTTTGAGAGG3'
NTR1F2	5'AGGATTCCAGTGGGACCTTTAAGCCAATTTATTTATGGGCCCTTGACCTCGGT CGACGGATCCCCGGGTT 3'
Gno1 F2	5'CGTTGATGGACTCCAAGGCACTGAATGAGATCTTTATGATAACAAACGACGGT CGACGGATCCCCGGGT 3'
Gno1 R1	5'TCCCCACTCTATATATCTTGCTGCGTGTGCAGACTGGCCAGCTGCTCACATCTG GAGGAAGTTTGAGAGG 3'
PFA1F2	5'TTTTTGCCAAAATTAAAAAGAATAGATCGGGTTTAAGACACAGTGAAAGTGG TCGACGGATCCCCGGGTT 3'
Pfa1 R1	5'AAAATAAAAAAATATGTATATACGCAATTGTTAAGAAAATGAGGATTAAATC TGGAGGAAGTTTGAGAGG 3'

Name	Sequence
Pfa1 R1 Longtine	5’AAAATAAAAAAATATGTATATACGCAATTGTTAAGAAAATGAATTCGAGCTC GTTTAAAC3’
Ntr1 R1 Longtine	5’TTTGTTTTTCGACAATAATATATAAAATCGTGCCTATCTCAGAATTCGAGCTCG TTTAAAC 3’
Gno1 R1 Longtine	5’TATATATCTTGCTCGTGTGCAGACTGGCCAGCTGCTCACGAATTCGAGCTCGT TTAAAC 3’
E-C2	5’ GAAAAGGCCAGCAATTTCAAGTTAACTCCAAAGAGTATCACTCAC 3’
D-A2	5’ AAACTTACAAGCCTAGCAAGACCGCGCACTTAAGCGCAGGCCCGG 3’
scR1	5’ CGTGTCTAGCCGCGAGGAAGGATTTGTTCC 3’
U3 PE	5' CCAAGTTGGATTCAGTGGCTC 3'
U1	5' GAATGGAAACGTCAGCAAACA 3'
NTR1-R3	5’AAAAGAACTTTTATCTGTGTTGGAGTCCGAATCCTCCATGCACTGAGCAGCGT AATCTG 3’
PFA1-R3	5’TACGTCTGCTTCCTTGATAATGACTATGCCTTTTTGCCATGCACTGAGCAGCGT AATCTG 3’
GNO1-R3	5’AACCAAACCGCTGTTTGGTTCTTGTAGCTGCCAAACCCATGCACTGAGCAGCG TAATCTG 3’

Name	Sequence
NTR1-F4	5’TTCGAAATAGTACAACCGAGAGAGGTCGAAGAACTTAAGCGAATTCGAGCTC GTTTAAAC 3’
PFA1-F4	5’AGAAGTGATTTAATCGCAGAGTCAAATATAAGCATAAATTGAATTCGAGCTC GTTAAAC 3’
GNO1-F4	5’GAGATGAGTACTCAATAGTAACATATAGGCAGCTTACACCGAATTCGAGCTC GTTTAAAC3’
35S	5’CGCTGCTCACC AATGG3’
GNO1 PCR3	5’TTAGCGATGTTATGCTTTCCAC 3’
GNO1 PCR5	5’TAGCCTTGCCCACTCTCTTTCTC3’
GNO1 W38SBtm	5’CCATACCGGGTTTCGATCCAAACTTTTC 3'
GNO1 W38STop	5’GAAAAGTTTGGATCG AACCCGGTATGG 3'

3.2.3 Whole cell splicing extract sedimentation through glycerol gradients

Extracts were prepared as described in Chapter 2.2.5 with an additional 100,000 x *g* centrifugation clarification for 1 hour and sedimented through 10-30% glycerol gradients as described in Chapter 2.2.9. Northern blotting for snRNAs was performed as in Chapter 2.2.9.

3.2.4 GFP Coimmunoprecipitations

GFP coimmunoprecipitations were performed with either with Abcam anti-GFP antibody (ab1218) or Chromotek-GFP-TRAP beads (Allele Biotech, ACT-CM-GFA0050) as described in Chapter 2.2.5. For immunoprecipitations followed by primer extension analysis, RNAs were removed from resin by extraction with 1:1 volume of phenol-chloroform to wash buffer.

3.2.5 rRNA analysis

Whole cell RNA was prepared and transferred for northern blotting as in Chapter 2.2.6. Oligonucleotide sequences used for northern blotting were previously described (Combs *et al*. 2006) and (Table 3.2).

3.2.6 Pulse chase analysis

For *HIS3::GAL::HA-NTR1*, cultures were grown to an OD (A_{600}) of 0.3-0.4 at 31°C in S-MET medium supplemented with 2% galactose and 1% sucrose. Cells were shifted to media containing 2% dextrose for 3 hours, and a low OD (A_{600}) was maintained throughout. For *prp2-1*, cultures were grown to an OD (A_{600}) of 0.3-0.4 at 31°C in YPD then shifted to YPD media pre-warmed to NPT. *prp2-1* cultures were grown at 37°C for 1 hour before harvesting 40 OD (A_{600}) units for pulse chase as in Chapter 2.2.7.

3.2.7 Native gel analysis

Extracts were prepared as described above; however, further clarification and buffer dialysis was performed as in Chapter 2.2.9. Native gel analysis and U6 snRNA northern blotting was completed as described in Chapter 2.2.9.

3.2.8 Fluorescence microscopy

Cultures were shifted from non-repressive media to SD-URA media supplemented with 2% galactose and 1% sucrose for 9 hours. Formaldehyde was added to 10% and samples incubated at 25°C for 30 min. Cells were pelleted and washed once with 1X PBS. Pellets were then resuspended in KSORB (1.1M sorbitol, 0.1 M KPO_4, 0.05% TritonX-100, pH 6.6) and incubated at room temperature for 6 minutes. Cells were washed once more with 1X PBS and resuspended in Vectashield mounting medium containing DAPI (Vector Labs, #H-1500) prior to fluorescence microscopy. Images were captured using a Leica DM IRBE microscope with a Leica DFC350 FX fluorescence camera. Merged images were processed in Photoshop.

3.2.9 MYC immunoprecipitations

SS3211 (*GAL::HA-GNO1 PRP43-GFP pRS416-GNO1-myc*) and SS3212 (*GAL::HA-GNO1, PRP43-GFP,* pRS416-*gno1 W38S-myc)* were grown in 1% galactose and 2% sucrose until OD (A_{600}) 0.4 then shifted to medium supplemented 2% dextrose for 6 hours. Extracts were prepared as described in 2.2.5 and incubated with anti-myc conjugated agarose (Sigma, #A7470) for 2 hours at 4°C with rotation. Resin was washed with IPP150 (10 mM Tris, pH 8.0, 150 mM NaCl, 0.1% NP-40, 1.5 mM $MgCl_2$, 8% glycerol, 1 mM DTT) 4X at 2500 x *g*. Bound proteins and RNAs were eluted by adding 1:1 phenol-chloroform to beads resuspended in wash buffer.

3.3 RESULTS

3.3.1 A dominant negative *prp43* mutant exhibits a first step splicing defect similar to the depletion of Ntr1p

Ntr1p has been shown to be necessary for proper snRNP recycling, and its absence results in a first-step splicing defect (Boon *et al*. 2006). Using our *GAL::HA-NTR1* depletion strain, we verified that U3 pre-snoRNA transcripts accumulate when this disassembly factor is not present (Fig. 3.1A). If this phenotype were dependent upon the stimulation of Prp43p by Ntr1p, it would be predicted that mutations in the ATPase domain of Prp43p would also result in a first-step splicing defect. Our laboratory has previously demonstrated that U3 pre-snoRNAs do not accumulate in two cold sensitive *PRP43* ATPase mutants (Combs *et al*. 2006). However, other alleles of *PRP43*, such as G395A and the cold sensitive dominant negative mutant Q423N, have been shown to accumulate introns (Arenas and Abelson 1997; Leeds *et al*. 2006). We analyzed U3 pre-snoRNAs after shifting the Q423N mutant from PT to NPT. U3 pre-snoRNA levels did increase at NPT (Fig. 3.1B), indicating the lack of Prp43p helicase activity manifests in defects at an earlier step in pre-mRNA splicing.

The steady state snRNP distribution is also disrupted when Ntr1p is depleted, due to a deficiency in the amount of the U5 snRNP (Boon *et al*. 2006). We analyzed whole cell snRNP distribution in the prp43Q423Np mutant at NPT by separating whole cell splicing extract though 10-30% glycerol gradients. The snRNAs present in each fraction of the gradient were detected by northern blotting (Fig. 3.1C & 3.1D). Similar to the depletion of Ntr1p, the distribution of the U4 and U6 snRNAs in spliceosomal particles

was perturbed in the Q423N mutant compared to wild-type (Boon *et al*. 2006; cf. Figs. 3.1C & 3.1D). In the Q423N mutant, the U4 snRNA sedimented higher in the gradient (fractions 10-14) compared to wild-type (fractions 16-22). These data suggest that the U4 snRNP is not being incorporated into larger splicing complexes. Additionally, we noted that the amount of the free U6 snRNP, which runs in fractions 8-12 in the wild-type, was reduced in the Q423N mutant. The total amount of the U5 snRNP was reduced by 30% in the Q423N mutant and the sedimented deeper in the gradient (Fig. 3.1E), which phenocopies the depletion of Ntr1p (Boon *et al*. 2006). We also found the same perturbation in snRNP distribution when the G-patch cofactor Pfa1p was overexpressed (Fig. 2.11).

3.3.2 Overexpression of Gno1p also affects the first-step of pre-mRNA splicing

If Gno1p is capable of competing with Ntr1p for Prp43p occupancy, it would be expected that the overexpression of Gno1p could indirectly inhibit pre-mRNA splicing. As in Figure 3.1A, we analyzed total RNA harvested from different time points throughout a time course in which Gno1p-HA/ZZ is overexpressed. After growth in non-repressive media, cells containing the overexpression vector p*GAL-GNO1-HA/ZZ* were shifted to media supplemented with 2% galactose and 1% sucrose. In the presence of overexpressed Gno1p-HA/ZZ, U3 pre-snoRNAs accumulated to a modest level (Fig. 3.2A), indicative of a mild pre-mRNA splicing defect. The U3 pre-snoRNA levels were only approximately two fold higher when Gno1p-HA/ZZ was present, (Fig. 3.2B), which was more modest than the effect of Ntr1p depletion. This intermediate effect on pre-

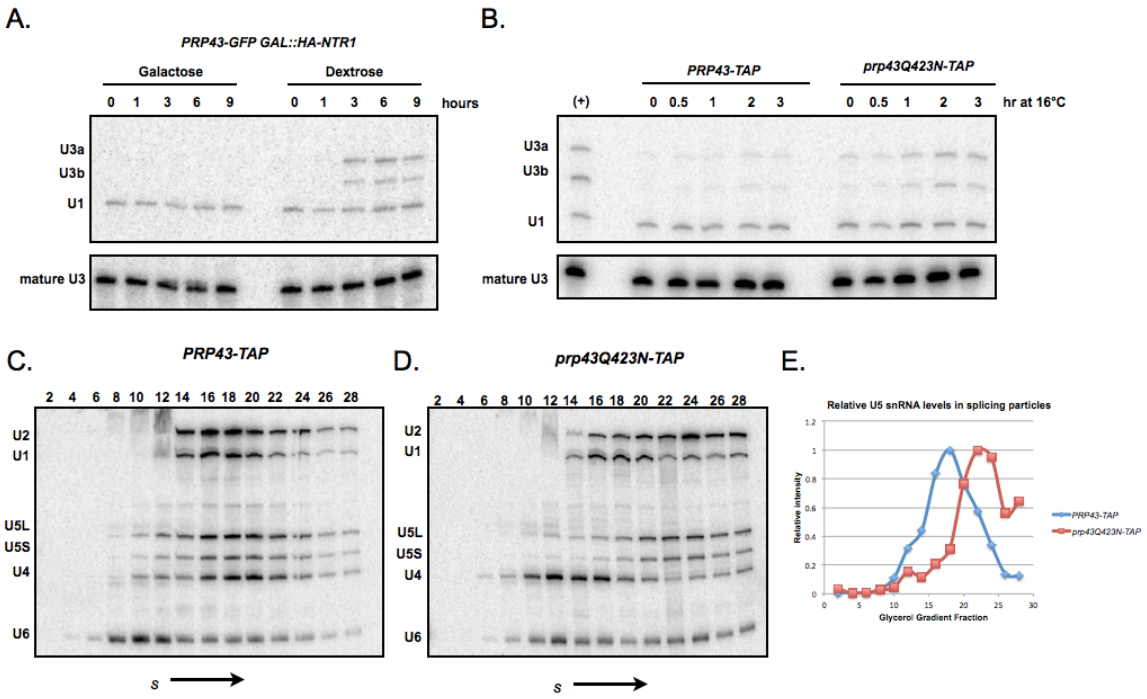


Figure 3.1 snRNP distributions are abrogated in *prp43Q423N*.

A). U3 pre-snoRNAs accumulated as Ntr1p was depleted. Primer extension analysis for U3 pre-snoRNAs were conducted on total RNA throughout a 9 hour time course depleting HA-Ntr1p. The U1 snRNA has been included as an internal control.
B). U3 pre-snoRNAs accumulated in the Q423N mutant when shifted to NPT. Total RNA was harvested at the time points indicated after shifting from 31°C to 16°C. Primer extensions same as (A). The positive control shown is from the 6 hour time point in (A).
C-D). Splicing extracts were prepared either wild-type *PRP43-TAP* (C) or *prp43Q423N-TAP* (D) after shift to 16°C for 3 hours. Glycerol gradients and northern blotting are presented as in Figure 2.11.
E). Quantitation of the U5 snRNA revealed decreased amount of mono-U5 snRNP.

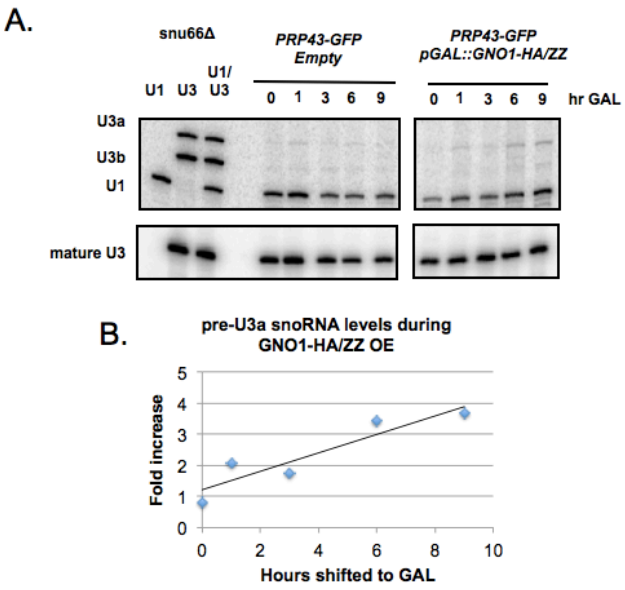


Figure 3.2 Gno1p overexpression impacts pre-mRNA splicing.

A). A mild splicing defect presents during the overexpression of Gno1p-HA/ZZ. Primer extension analysis for the U3 pre-snoRNA from total RNA prepared at the indicated time points after shift to S-URA media supplemented with 1% sucrose and 2% galactose. RNA prepared from a *snu66Δ* strain grown at 16°C for 6 hours was used as a positive control for pre-snoRNA accumulation (Stevens *et al*. 2001).
B). Quantitation for the relative fold increase of pre-U3a signals in (A) and trendline is graphed. U3a levels were normalized to corresponding wild-type time points to account for the influence of shifting to a different carbon source.

mRNA would be expected if Gno1p does partially compete with Ntr1p for Prp43p binding (see below).

3.3.3 Loss of Prp43p stimulation affects tri-snRNP assembly

We previously observed a defect in the assembly of tri-snRNP in the presence of overexpressed Pfa1p-HA/ZZ (Fig. 2.12), and proposed a model in which excess Pfa1p inhibits Ntr1p binding and stimulation of Prp43p so that proper snRNP assembly is indirectly affected (Illustration 2.1). The loss of Ntr1p binding or a mutation in the ATPase domain of Prp43p would be expected to produce similar tri-snRNP formation defects. As described above, we prepared whole cell splicing extracts from *GAL::HA-NTR1* after depleting HA-Ntr1p for 6 hours and from the *prp43Q423N-TAP* mutant strain after growth at NPT. Splicing extracts were separated through a native gel and complexes detected by northern blotting. As predicted, the steady-state snRNP levels shifted from tri-snRNP to di-snRNP when HA-Ntr1p is depleted (Fig. 3.3A, lanes 1 & 2; Boon *et al*. 2006). We also show a strong decrease in U4/U6•U5 levels in the Q423N at 16˚C (Fig. 3.3A, cf. lanes 7 & 8). The U4/U6 di-snRNP levels were increased by over 6-fold in the Q423N compared to the wild-type at NPT (Fig. 3.3B). The amount of tri-snRNP was reduced by half in the *prp43* ATPase mutant. Additionally, the U6 snRNP levels were decreased in *prp43Q423N* relative to wild-type, as well as in the depletion of HA-Ntr1p or the overexpression of Pfa1p-HA/ZZ. Therefore, Prp43p stimulation appears to be required for proper snRNP equilibrium and spliceosome assembly, most likely due to a lack of U5 recycling (Boon *et al*. 2006).

Overexpression of Gno1p-HA/ZZ results in a mild accumulation of U4/U6 di-snRNP (Fig. 3.3, cf. lanes 5 & 6). While U4/U6 di-snRNP levels increase to over 2-fold that of wild-type, there is little measurable decrease in tri-snRNP (Fig. 3.3B). This intermediate effect on tri-snRNP assembly correlates to the moderate effect on pre-mRNA splicing in the presence of excess Gno1p (Fig. 3.2). These data further suggest that overexpression of Gno1p can compete with Ntr1p for Prp43p occupancy.

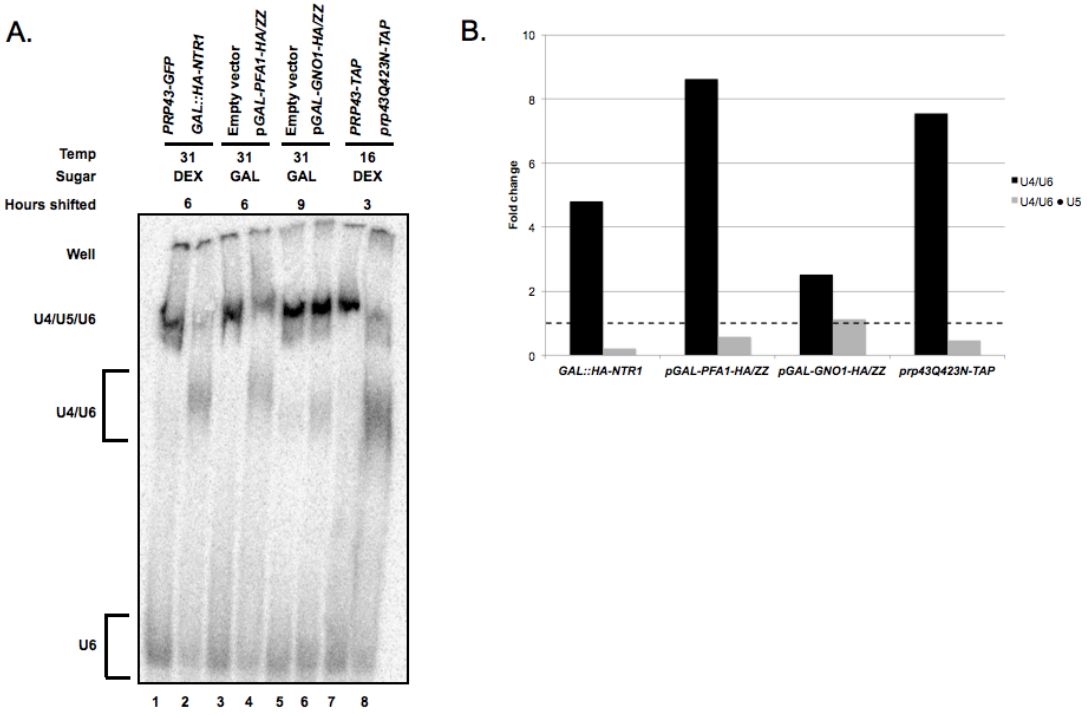


Figure 3.3 Fluctuation in G-patch protein levels result in pre-mRNA splicing defects and affects tri-snRNP assembly.

A). Native gel northern blot analysis of the U4/U6 and U4/U6•U5 defects observed when G-patch protein levels are lowered or increased. Splicing extracts were prepared from strains under the conditions indicated, and 50 µg of total protein were separated on a 4% [80:1] acrylamide native gel followed by northern blot analysis for the U6 snRNA.
B). Graphical representation of the data from (A). The graph represents the fold increase or decrease in U4/U6 (black bars) and U4/U6•U5 (grey bars) seen in each of the conditions relative to its corresponding WT sample. The dashed black line represents the wild-type baseline.

3.3.4 Prp43p subcellular localization is affected by overexpression of Gno1p or Ntr1p

Gno1p and Ntr1p localize to the nucleolus and the nucleoplasm, respectively (Guglielmi and Werner 2002; Huh *et al*. 2003). If Gno1p and Ntr1p compete for Prp43p binding, the localization of Prp43p-GFP should shift from one subcellular compartment to the other. Under wild-type circumstances, Prp43p localized primarily to the nucleolus (Combs *et al*. 2006), but a nucleoplasmic signal was evident was well (Fig. 3.4A & 3.4B, first row). We analyzed Prp43p-GFP localization in the presence of overexpressed Gno1p-HA/ZZ, Ntr1p-HA/ZZ or Pfa1p-HA/ZZ. Interestingly, Prp43p-GFP localization to the nucleoplasm was not observed when Gno1p-HA/ZZ was in excess (Fig. 3.4A, second row). In the majority of the cells analyzed, a clear separation was apparent between the DAPI stained DNA and the Prp43p-GFP signal (Fig. 3.4B, second row) indicating increased partitioning to the nucleolus. In contrast, if Ntr1p-HA/ZZ was overexpressed, Prp43p-GFP localization to the characteristic nucleoli caps in the nucleus was disrupted (Fig. 3.4A, third row). The Prp43p-GFP was more diffuse within the nucleoplasm under these conditions, and very little difference between the GFP and DAPI signals was observed (Fig. 3.4B, third row). Unlike either the overexpression of Ntr1p-HA/ZZ or Gno1p-HA/ZZ, increased levels of Pfa1p-HA/ZZ did not significantly affect the subcellular localization of Prp43p-GFP (Fig. 3.4A, last row). Perhaps since Pfa1p localizes to the nucleoplasm and is predicted to reside in the nucleolus (Huh *et al*. 2003; Lebaron *et al*. 2009), no overall change in Prp43p-GFP localization occurred even though Ntr1p and Gno1p associations are disrupted (Fig. 2.8). However, these data

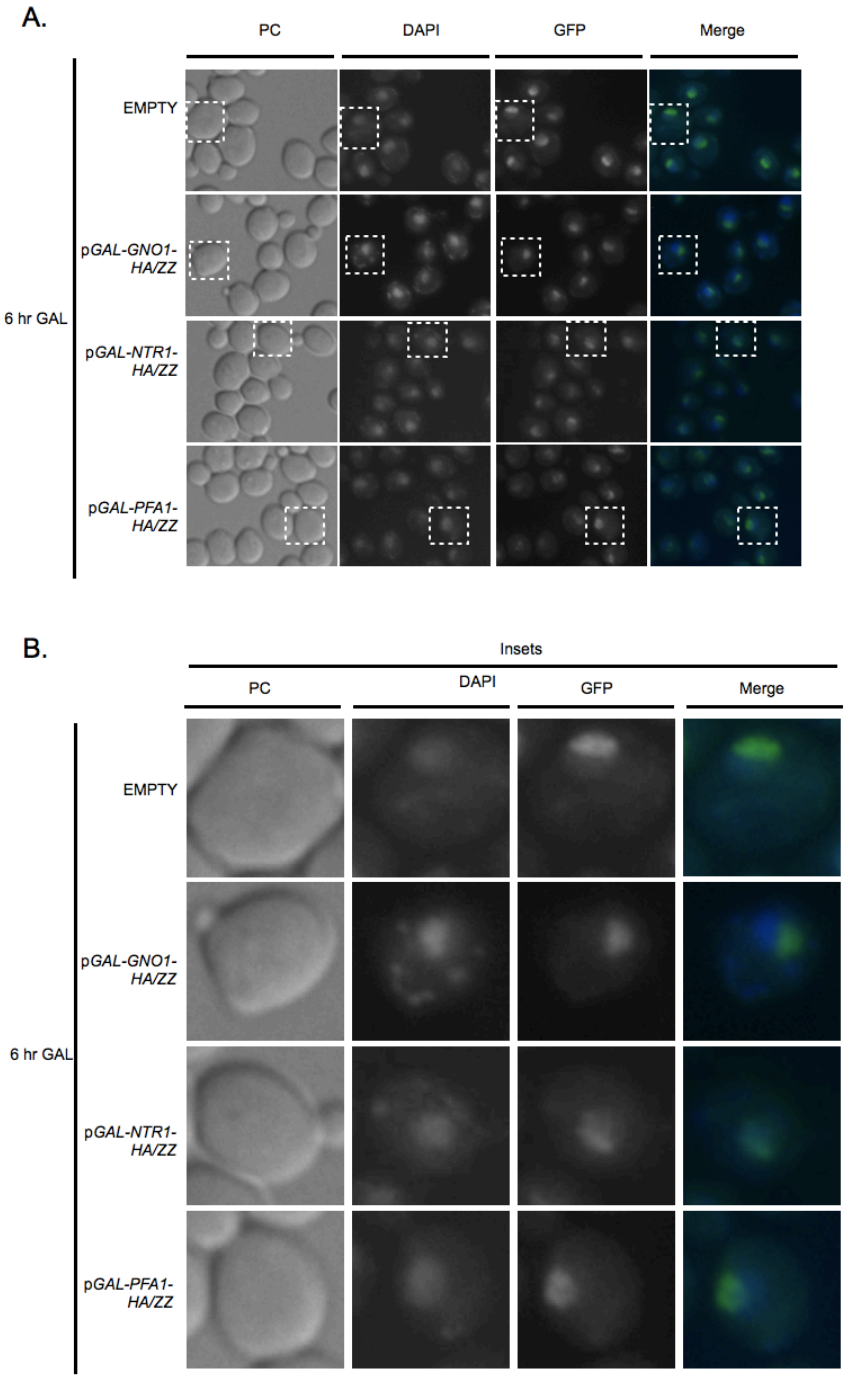


Figure 3.4 Prp43-GFP delocalizes upon cofactor overexpression.

(Fig. 3.4 contd.)

A). Fluorescent microscopy visualization of Prp43p-GFP after shift to media containing 2% galactose and 1% sucrose for 6 hours in the presence of either empty vector (top row), p*GAL-GNO1-HA/ZZ* (second row), p*GAL-NTR1-HA/ZZ* (third row), or p*GAL-PFA1-HA/ZZ* (fourth row). Fields for phase contrast (PC) (first column), DAPI [4',6'-diamidino-2-phenylindole] (second column) and GFP (third column) channels are shown. DAPI and GFP channels are merged in fourth column. Dashed boxes indicated cells presented at higher magnification in (B). B). Selected cells are shown at higher magnification to highlight nucleoplasmic versus nucleolar localization.

further suggest that Ntr1p and Gno1p have the ability to compete with each other for Prp43p since changes in their steady state levels impact Prp43p localization.

3.3.5 Increased levels of Gno1p can preclude Ntr1p association with Prp43p

The moderate tri-snRNP assembly defect (Fig. 3.3) along with the shift in Prp43p-GFP localization away from the nucleoplasm (Fig. 3.4) indicates that Gno1p-HA/ZZ can sequester Prp43p from Ntr1p and the pre-mRNA splicing machinery. By performing coimmunoprecipitations against Prp43p-GFP in the presence of wild-type or overexpressed levels of Gno1p, we analyzed Prp43p-GFP/Ntr1p-myc association. Strains carrying either an empty vector or p*GAL-GNO1-HA/ZZ* plasmid were shifted to media supplemented with 2% galactose 1% sucrose for 6 hours. Subsequent western blot analysis of the Prp43p-GFP purified material for Ntr1p-myc indicated that under wild-type conditions, Ntr1p still associated with Prp43p-GFP (Fig. 3.5, lane 12). In contrast, when Gno1p-HA/ZZ was overexpressed, Ntr1p-myc/Prp43p-GFP complexes were reduced (Fig. 3.5, lane 15). These data indicate that the pre-mRNA splicing defects observed when Gno1p-HA/ZZ is overexpressed are a result of the loss of interaction between Prp43p-GFP and Ntr1p-myc. We note that little Gno1p-HA/ZZ actually associated with Prp43p-GFP (WB PAP, lane 15). Perhaps a more complex sequestration of Prp43p-GFP in the nucleolus independent of Gno1p is at work downstream of the competition between Gno1p and Ntr1p.

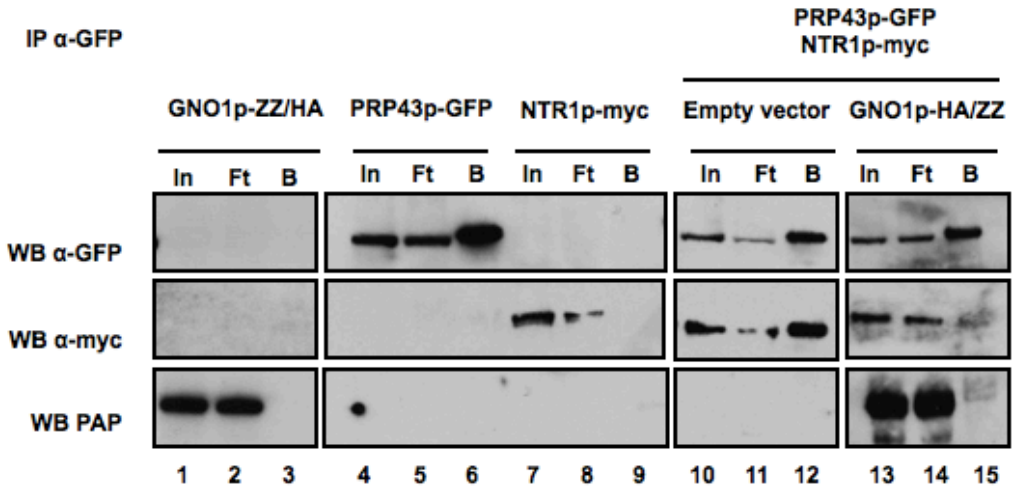


Figure 3.5 Overexpression of Gno1p reduces the levels of Prp43p-GFP/Ntr1p complexes.

Strains harboring either an empty vector (pRS316-*GAL*) or p*GAL-GNO1-HA/ZZ* were shifted to media supplemented with 2% galactose and 1% sucrose for 6 hours, then Prp43p-GFP complexes were purified using α-GFP antibodies. Anti-GFP western blots in the top panels demonstrate Prp43p-GFP IP efficiency. The anti-myc western blots in the second row show Ntr1p-myc signals. Ntr1p-myc is present in Prp43p-GFP purified material in the presence of the empty vector (lane 12), but a reduced amount is present when Gno1p-HA/ZZ is overexpressed (lane 15). Bottom panels demonstrate the overexpression of Gno1p-HA/ZZ.
IP, immunoprecipitated; WB, western blot; IN, 1% input; FT, 1% flow-through; B, 20% bound

3.3.6 A mutation in *GNO1* abrogates Gno1p and Prp43p association with 35S rRNA

The study initially characterizing Gno1p as a ribosome biogenesis factor also showed that its G-patch domain is critical for its function (Guglielmi and Werner 2002). A single mutation in a conserved tryptophan of the G-patch domain (W38S) resulted in a growth defect similar to *gno1*Δ (Guglielmi and Werner 2002). Since the G-patch domain of Ntr1p is also necessary for its interaction with Prp43p (Tanaka *et al*. 2007), we hypothesized that the G-patch domain in Gno1p may have a similar role. We created centromeric plasmids harboring either myc-tagged wild-type *GNO1* or *gno1W38S*, under the control of its own promoter and introduced them into the *GAL::HA-GNO1* depletion strain. While the plasmid-borne *GNO1* rescued the growth of *GAL::HA-GNO1* in the presence of dextrose, *gno1W38S* did not (Fig. 3.6A). Using these strains, we depleted the HA-Gno1p and analyzed the association of the wild-type or mutant Gno1p-myc with Prp43p-GFP. We immunoprecipitated Prp43p-GFP and western blotted for the presence of wild-type or mutant Gno1p-myc. We also performed the reverse immunoprecipitation using the MYC tag on Gno1p and western blotted for the presence of Prp43p-GFP. In both cases, Prp43p-GFP/Gno1p-myc association was reduced by 60% in the G-patch mutant (Fig. 3.6B). These data indicate that the G-patch domain of Gno1p is important for its interaction with Prp43p *in vivo*.

We then analyzed the ability of gno1W38Sp-myc to associate with the 35S pre-rRNA since Prp43p and Gno1p exist together in 90S particles. After immunoprecipitating with wild-type Gno1p or the W38S mutant, we used primer extension analysis to determine the amount of 35S pre-rRNA in the same purified material as in Fig. 3.6B. We

found that the W38S mutation in Gno1p decreased its association with the 35S pre-rRNA by 80% (Fig. 3.6C, right bars). We performed a similar experiment to determine if Prp43p/35S pre-rRNA association was also abrogated in the *gno1W38S* mutant. Consistent with a function for Gno1p in recruiting Prp43p to the 35S transcript, we found that the amount of pre-rRNA associated with Prp43p-GFP was also decreased in the W38S mutant (Fig. 3.6C, left bars). A previous study had indicated that Prp43p association with the pre-35S rRNA was not dependent on Gno1p (Lebaron *et al*. 2009). However, we consistently found that after depletion of HA-Gno1p the amount of Prp43p-GFP associated with the 35S pre-rRNA was reduced (Fig. 3.6D). This is likely due to the fact that we have transiently depleted Gno1p in our experiments, and they use a *gno1*Δ strain has additional collateral *PFA1* and *LTV1* defects which could provide suppressor phenotypes to this phenomenon (Lebaron *et al*. 2009). Neither the depletion of Gno1p nor the W38S mutation resulted in complete loss of Prp43p-GFP association with the 35S pre-rRNA, but our data indicate that Gno1p greatly enhances its incorporation into 90S particles. Furthermore, the G-patch domain of Gno1p is important in these interactions, making it similar to that of Ntr1p (Tanaka *et al*. 2007).

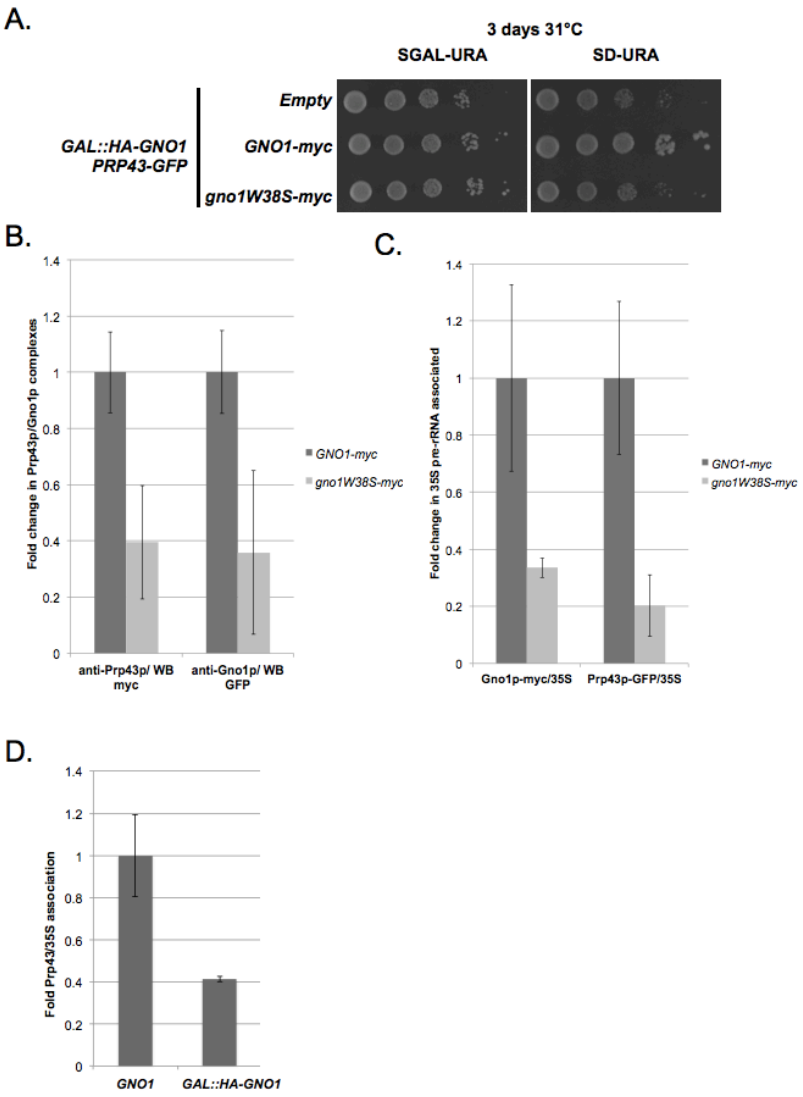


Figure 3.6 The *gno1W38S* mutant is impaired for Prp43p and 35S pre-rRNA association.

(Fig. 3.6 contd.)

A. Growth assays of *GAL::HA-GNO1* depletion strains harboring empty vector (pRS416), pRS416-*GNO1-MYC* or pRS416-*gno1W38S-MYC* on media supplemented with either galactose or dextrose. Plates were photographed after 3 days growth at 31˚C. B. Gno1p-myc and Prp43p-GFP associate with less 35S pre-rRNA in the *gno1W38S*
(Fig 3.6 contd.) mutant. Quantitation of 35S pre-rRNA signal associated with Gno1p-myc or Prp43p-GFP after depletion of HA-Gno1p for 6 hours is shown. Gno1p-myc wild-type (dark grey) or gno1W38Sp-myc (light grey) associated material was immunoprecipitated and the presence of 35S pre-rRNA determined by primer extension. The averages and standard deviations of 35S pre-rRNA association in biological replicates (n=3) are graphed (p-value= 0.025). Similarly, from the same extracts, Prp43p-GFP was immunoprecipitated and 35S pre-rRNA association determined by primer extension (light gray bars); p-value= 0.0087. C. Prp43p/Gno1p association is disrupted in the *gno1W38S* mutant. From the immunoprecipitations described in (B), western blot analysis was performed for Prp43p-GFP or Gno1p-myc. Anti-myc western blots were performed to detect Gno1p-myc (dark grey) or gno1W38Sp-myc (light grey) contained in Prp43p-GFP purified material. Averages and standard deviations are graphed (n=3; p-value=0.027). The reverse immunoprecipitations for Prp43p-GFP from the same extracts were performed followed by western blot for Gno1p-myc or gno1W38Sp-myc (right bars) (n=3; p-value= 0.013). D. Prp43p consistently associates with a reduced amount of 35S pre-rRNA in the absence of Gno1p. The 35S pre-rRNA immunoprecipitated with Prp43p-GFP after HA-Gno1p depletion was analyzed by primer extension as described in (C).

3.3.7 Loss of Ntr1p indirectly affects ribosome biogenesis

Through a screen of ts alleles of essential genes, *NTR2* was determined to have an effect on ribosome biogenesis. The *NTR2* ts allele had increased steady state levels of the 35S pre-rRNA, as well as the 23S aberrant species. Concomitantly, both the 20S and 27S pre-rRNA species were decreased (Ben-Aroya *et al*. 2008). Since Ntr1p is found primarily in a complex with Ntr2p (Tsai *et al*. 2005), similar ribosome biogenesis defects might be expected in cells deficient for Ntr1p function. We show that as HA-Ntr1p is depleted the 35S pre-rRNA and the 23S aberrant species accumulated, and the 20S pre-rRNA levels concomitantly decreased (Fig. 3.7A). Additionally, the 27S levels increased until all of the rRNA signals drastically decreased at the time of cell death, approximately 9 hours after shift to dextrose (Fig. 3.7A; data not shown). Indeed, loss of Ntr2p and Ntr1p function results in similar ribosome biogenesis defects.

We further explored the pre-rRNA processing defect in the *GAL::HA-NTR1* depletion strain by pulse chase analysis. After depletion of HA-Ntr1p for 3 hours, we observed that the 35S precursor persisted until 2 minutes after the chase (Fig. 3.7B). The kinetics of 27S rRNA processing was slowed as well, as evidenced by its presence 5 minutes after the chase. The amount of 20S pre-rRNA at time 0 was decreased, consistent with the northern blot results.

Although Ntr1p is a factor involved in spliceosome disassembly, its interaction and stimulation of Prp43p is necessary for proper spliceosome recycling and the first step of pre-mRNA splicing (Fig. 3.1 & 3.3). Historically, mutants of pre-mRNA splicing factors that affect the first step of splicing also impact ribosome biogenesis (Warner and

Udem 1972; Shulman and Warner 1978). To determine how the rRNA processing defects observed when HA-Ntr1p is depleted compares to that of a spliceosome assembly mutant, we analyzed rRNA processing by northern blotting and pulse chase analysis in the ts *prp2-1* mutant. A shift to NPT inactivates Prp2p in this strain so that it does not join the spliceosome to make the essential rearrangements for the first transesterification reaction (King and Beggs 1990). Northern blotting for the rRNA species in *prp2-1* at NPT we observed that the 35S pre-rRNA accumulated, while the levels of 27S and 20S pre-rRNA species decreased, as in *GAL::HA-NTR1* (Fig. 3.7C). The 23S aberrant species is not as evident by northern blot, but is observed by pulse chase analysis in *prp2-1* after a 1 hours shift to NPT (Fig. 3.7D). These data suggest that the loss of Ntr1p results in similar rRNA processing defects as other spliceosomal factors; therefore, these defects are most likely to be indirect and due to the loss of r-proteins (Shulman and Warner 1978).

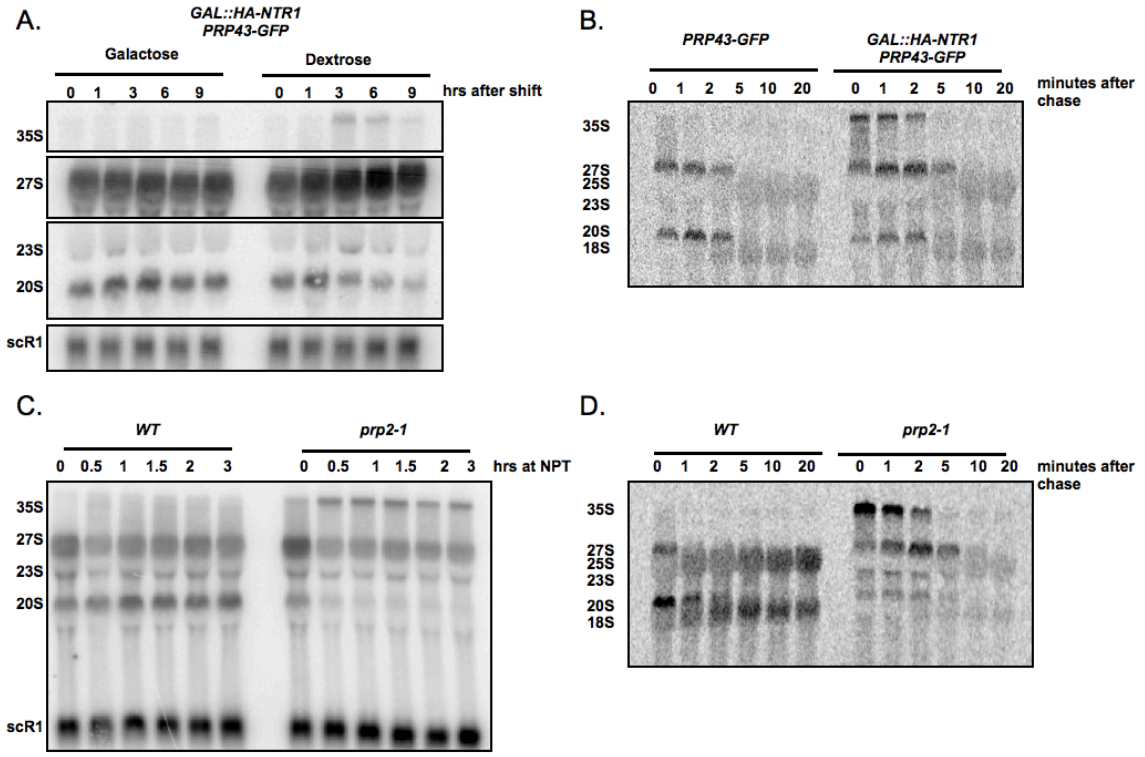


Figure 3.7 Loss of Prp2p and Ntr1p result in similar ribosome biogenesis defects.

A). Loss of Ntr1p results in accumulation of the 35S pre-rRNA and eventual cell death. Northern blot analysis of pre-rRNA intermediates in the presence of Ntr1p (lanes 1-5) and in cells where HA-Ntr1p is depleted over a time course (lanes 6-10). Increases in the 35S pre-rRNA and the aberrant 23S intermediate, along with a concomitant decrease in 20S pre-rRNA are observed until 9 hours at which time cells are likely to be dying. The scR1 RNA loading control is shown at the bottom of the panel. B). Pulse chase analysis of rRNA intermediates after metabolic depletion of HA-Ntr1p for 3 hours. Cells containing Ntr1p show normal rRNA processing kinetics after chase with cold methionine (lanes 1-6). After Ntr1p has been depleted for 3 hours, delayed kinetics of 35S, 27S and 20S rRNAs are evident (lanes 7-12). C). Loss of Prp2p function correlates with a 35S rRNA accumulation and a severe 20S depletion phenotype. Northern blot analysis of pre-rRNA intermediates in RNA harvested from *PRP2* cells (lanes 1-5) is shown. RNA harvested from *prp2-1* cells after growth at the NPT for the indicated times are shown in lanes 6-10. Steady state accumulation of the 35S pre-rRNA and rapid decrease in the 20S pre-rRNA levels are noted. The scR1 RNA loading control is shown at the bottom of the panel. D). Pulse chase analysis of RNA harvested from *prp2-1* cells after shift to non-permissive temperature for 1 hour is shown.

3.3.8 Overexpression of Ntr1p-HA/ZZ impacts rRNA processing and Prp43p binding with Gno1p-TAP

If Ntr1p and Gno1p can directly compete for Prp43p binding, then overexpression of Ntr1p-HA/ZZ would be expected to perturb ribosome biogenesis. We analyzed steady state rRNAs by northern blot from a time course in which Ntr1p-HA/ZZ was overexpressed. In the presence of excess Ntr1p-HA/ZZ there was a mild increase in the steady state levels of the 35S pre-rRNA, but there was no reciprocal decrease in the 27S and 20S pre-RNAs (Fig. 3.8A). These results are milder but similar to what we observed in our p*GAL-PFA1-HA/ZZ* overexpression strain (Fig. 2.6).

To directly determine if Ntr1p-HA/ZZ overexpression can disrupt Gno1p/Prp43p binding, we performed coimmunoprecipitations against Prp43p-GFP in the presence of wild-type or overexpressed levels of Ntr1p. Ntr1p-HA/ZZ was overexpressed by shifting strains to media supplemented with 2% galactose 1% sucrose for 6 hours. Western blot analysis for Gno1p-TAP in the Prp43p-GFP purified material indicates that Gno1p-TAP associated with Prp43p-GFP under wild-type conditions (Fig. 3.8, lane 12). However, when Ntr1p-HA/ZZ was overexpressed, Prp43p-GFP interacted with a reduced amount of Gno1p-TAP (Fig. 3.5, lane 15). Together with the shift of Prp43p-GFP localization and the rRNA processing defects, these data indicate that excess Ntr1p can compete with Gno1p for Prp43p occupancy. In both cases, Gno1p and Ntr1p exert milder effects over one another, in contrast to the dominant effect of overexpressed Pfa1p over the other two cofactors.

3.3.9 Pfa1p/Prp43p association is not disrupted by overexpression of Ntr1p or Gno1p

If Ntr1p and Gno1p can directly compete with Pfa1p for Prp43p occupancy, then overexpression of either Gno1p-HA/ZZ or Ntr1p-HA/ZZ would be expected to affect Prp43p association with Pfa1p. We performed coimmunoprecipitations against Prp43p-GFP in the presence of wild-type or overexpressed levels of Ntr1p and observed that Pfa1p still efficiently associated with Prp43p even when Ntr1p-HA/ZZ was overexpressed (Fig. 3.9, lane 24). Similarly, we found that the overexpression of Gno1p-HA/ZZ did not disrupt Prp43p/Pfa1p complexes (Fig. 3.9, lane 30). Therefore, neither the overexpression of Ntr1p nor Gno1p affects Prp43p association with Pfa1p, consistent with the strength of the interaction discussed above.

.

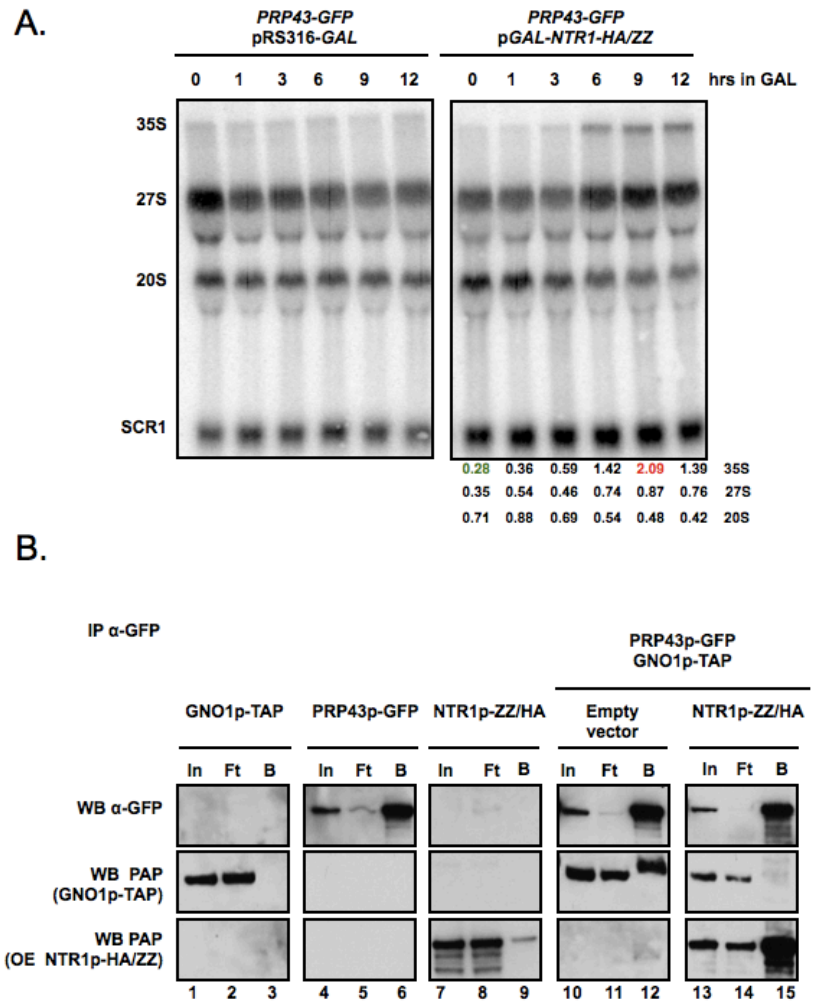


Figure 3.8 Ntr1p overexpression disrupts Gno1p/Prp43p association.

A). Northern blot analysis of rRNAs during a time course overexpressing Ntr1p-HA/ZZ shows a mild ribosome biogenesis defect. rRNA species are indicated to the left. Quantification of normalized signals for 35S, 27S, and 20S pre-rRNAs are shown below the panel. B). Prp43p-GFP association with Gno1p-TAP is disrupted in Ntr1p-HA/ZZ overexpression. Western blot analysis for Prp43p-GFP (top), Gno1p-TAP (middle), and Ntr1p-HA/ZZ (bottom) were performed following Prp43p-GFP immunoprecipitation from strains with wild-type (lanes 10-12) or overexpressed (lanes 13-15) levels of Ntr1p. IP, immunoprecipitated; WB, western blot; IN, 1% input; FT, 1% flow-through; B, 20% bound

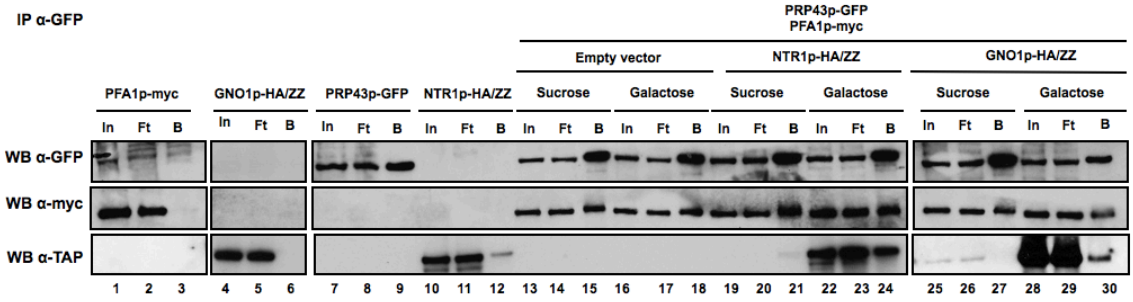


Figure 3.9 Neither overexpressed Ntr1p nor Gno1p disrupt Prp43p/Pfa1p complexes.

Strains harboring either an empty vector (pRS316-*GAL*), p*GAL-NTR1-HA/ZZ,* or p*GAL-GNO1-HA/ZZ* were shifted to media supplemented with 2% galactose and 1% sucrose for 6 hours, then Prp43p-GFP complexes were purified using α-GFP antibodies. Anti-GFP western blots in the top panels demonstrate the efficiency of the Prp43p-GFP immunoprecipitation. Pfa1p-myc association with Prp43p-GFP is shown in the middle panels (WB α-myc). Bottom panels demonstrate the overexpression of Ntr1p-HA/ZZ or Gno1p-HA/ZZ.

IP, immunoprecipitated; WB, western blot; IN, 1% input; FT, 1% flow-through; B, 20% bound

3.4 DISCUSSION

We have previously shown that Ntr1p, Gno1p, and Pfa1p interact with Prp43p in a mutually exclusive manner (Figure 2.1), and that the interaction of Pfap1p with Prp43p is dominant over that of Gno1p or Ntr1p (Figures 2.9 & 2.13). We further investigated the relationships between Ntr1p and Gno1p to construct a complete network of Prp43p cellular distribution by the G-patch proteins. For Prp43p to be regulated by these cofactors, each cofactor would be expected to be function in different steps of gene expression. Mass spectrometry, biochemical and genetic data indicate that Gno1p functions in 90S pre-ribosomal particles and is not present in mature 60S or 40S pre-ribosomal subunits (Guglielmi and Werner 2002; Lebaron *et al*. 2009). We show that Gno1p associates with the 35S pre-rRNA as expected (Fig. 3.6). To date, Gno1p has no known connection to pre-mRNA splicing.

Mass spectrometry data indicate that Ntr1p only associates with splicing factors (Lebaron *et al*. 2009), but the depletion of Ntr1p has a strong effect on both pre-mRNA splicing and ribosome biogenesis, suggesting that it may play two roles (Fig. 3.7). However, we demonstrate that Ntr1p depletion phenocopies the rRNA processing defects as well as the pre-mRNA splicing defect (Boon *et al*. 2006, Fig. 3.1A) of the temperature sensitive *prp2-1* pre-mRNA splicing mutant (Warner and Udem 1972, Fig. 3.7). Since the loss of Ntr1p leads to the same first-step splicing defect as *prp2-1*, the effect on ribosome biogenesis is likely results from the loss of r-proteins, as is the case for *prp2-1* (Warner and Udem 1972; Shulman and Warner 1978; Bromley *et al*. 1982). Our investigations of Pfa1p have shown that it is not necessary for the role of Prp43p in spliceosome

disassembly (Fig. 2.10B) nor for ribosome biogenesis (Pertschy *et al*. 2009). While Pfa1p associates with pre-20S rRNA, it does not associate with snRNAs (Lebaron *et al*. 2005), so while Pfa1p may play an accessory role in 40S subunit maturation, it is unlikely that it also plays a similar role in pre-mRNA splicing. Therefore, we believe each G-patch protein functions in a distinct step of ribosome biogenesis or pre-mRNA splicing, so that it can faithfully regulate the activity of Prp43p in the proper pathway.

We have previously shown that excess Pfa1p competes with Ntr1p for Prp43p binding so that snRNPs are not recycled properly and U4/U6•U5 tri-snRNP levels decrease (Fig. 2.12). Data from Beggs and colleagues demonstrated a role for Ntr1p in the recycling of U5 (Boon *et al*. 2006), and our results indicate that if HA-Ntr1p is depleted, cells become deficient in tri-snRNP (Fig. 3.2). Our data suggests that the stimulation of the ATPase activity of Prp43p is necessary for U5 recycling. Steady state snRNP levels in the cs mutant *prp43Q423N-TAP* at NPT mimic that of *GAL::HA-NTR1* (Figs. 3.1C & 3.1D; Boon *et al*. 2006). Consequently, the Q423N mutant accumulates increased levels of U4/U6 di-snRNP and decreased levels of tri-snRNP (Fig. 3.3). This inability to recycle the spliceosome *in vivo* is consistent with the defect in the first step of splicing that is observed in either the Q423N mutant or when Ntr1p is depleted (Fig. 3.1A & 3.1B). Moreover, the overexpression of Gno1p, which moderately impacts Prp43p/Ntr1p association (Fig. 3.5), also results in a mild effect on pre-mRNA splicing (Fig. 3.2). Once again, under these conditions, the association of Prp43p with Ntr1p is abrogated leading to a 2-fold increase in U4/U6 levels (Fig. 3.3). Therefore, disassembly of the U2/U6•U5 complex by Prp43p through Ntr1p-mediated interaction with the

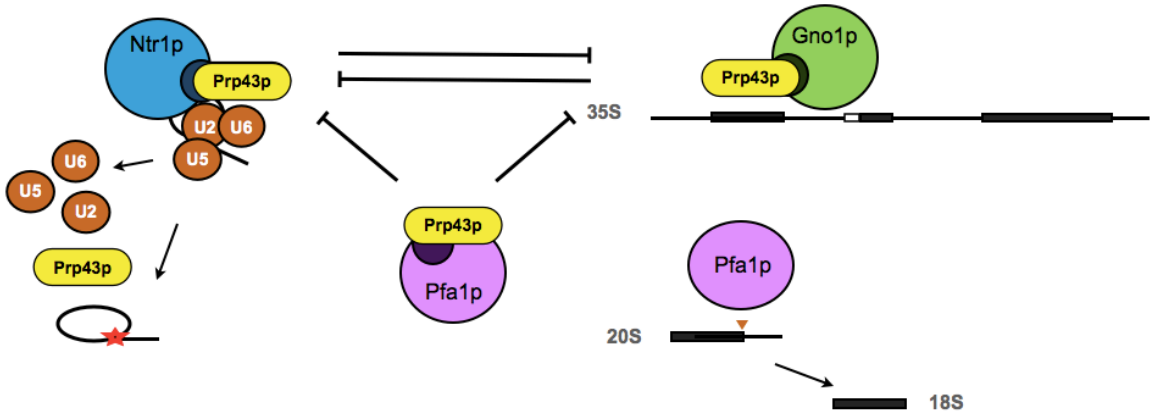


Illustration 3.1 Model of Prp43p regulation by G-patch protein cofactors.

Ntr1p and Gno1p can compete for Prp43p occupancy with each other, thereby moderating the incorporation of Prp43p in pre-mRNA splicing or pre-90S complexes. The G-patch domains of both Ntr1p and Gno1p are important in mediating their interactions with Prp43p. In contrast, the G-patch domain of Pfa1p is dispensable for its interaction with Prp43p. Additionally, Pfa1p can strongly inhibit the association of Prp43p with the other two cofactors, resulting in the pausing of both mRNA and rRNA processing pathways.

spliceosome is necessary for maintaining the proper equilibrium of snRNP levels. Such defects in disassembly quickly impact the processing of pre-mRNAs by the spliceosome *in vivo*.

Of the four G-patch proteins characterized in yeast, three have proven to play some role in recruitment and/or stimulation of RNA helicase activity. The G-patch domain of Spp2p is necessary for the incorporation of Prp2p into the spliceosome for the first step of splicing (Silverman *et al*. 2004). Similarly, the G-patch domain of Ntr1p is required for its interaction with Prp43p and via Ntr2p mediates incorporation into stalled spliceosomes or the post-splicing lariat RNP (Tsai *et al*. 2005; Tanaka *et al*. 2007). Likewise, Prp43p associates with less pre-20S in a *pfa1*Δ strain, indicating Pfa1p may facilitate Prp43p incorporation into pre-40S complexes (Lebaron *et al*. 2009). However, these G-patch cofactors appear to differentially impact Prp43p activity; the helicase activity of Prp43p has been shown to be stimulated by Ntr1p, and Pfa1p enhances both its helicase and ATPase activities (Tanaka *et al*. 2007; Walbott *et al*. 2010). It is yet to be determined whether Spp2p stimulates the activities of Prp2p.

While the G-patch domain of Pfa1p is not necessary for Prp43p binding, it is responsible for the stimulation of its activities (Pandit *et al*. 2009; Walbott *et al*. 2010). In contrast, the G-patch domains of Spp2p and Ntr1p are necessary to associate with Prp2p or Prp43p, respectively (Silverman *et al*. 2004; Tanaka *et al*. 2007). The G-patch domain of Gno1p is important for cell growth, like that of Ntr1p (Tanaka *et al*. 2007), and is necessary for its function in ribosome biogenesis (Guglielemi and Werner 2002). The recruitment of Prp43p to the 35S pre-rRNA was previously reported to not be strictly

dependent upon the presence of Gno1p (Lebaron *et al*. 2009), but Prp43p activity is stimulated by Gno1p *in vitro*, implying a functional regulatory relationship with Prp43p (Y. Henry, unpublished results). Our data demonstrate that the G-patch domain of Gno1p plays a role in its stable association with Prp43p. A single serine substitution at the conserved W38 in the G-patch domain reduced Gno1p/Prp43p complexes by 80% *in vivo* (Fig. 3.7B). Perhaps other secondary interfaces or proteins are present between Gno1p and Prp43p, but the G-patch domain appears to significantly stabilize their interaction. Furthermore, we show that the W38S mutant associated with less 35S pre-rRNA *in vivo*, and Prp43p also associated with less 35S pre-rRNA in the W38S mutant (Fig. 3.7C). While the association of neither Gno1p nor Prp43p with the pre-rRNA is completely lost, our data suggest that the G-patch domain functions in the stable association with the pre-rRNA. These data could support two separate models of Gno1p/Prp43p interaction. Gno1p could interact with Prp43p via its G-patch domain and together they incorporate into 90S complexes. Alternatively, Prp43p can associate with the 35S pre-rRNA independently, then Gno1p, which interacts via its G-patch domain with the rRNA, stabilizes this interaction. Further *in vitro* analysis of Gno1p and Prp43p will be necessary to elucidate the nature of this interaction. Nevertheless, the G-patch of Gno1p and Ntr1p are necessary for their Prp43p-mediated functions *in vivo*. We predict that these two cofactors work in a similar manner with Prp43, in contrast to Pfa1p. The ability of Ntr1p and Gno1p to compete with one another, while Pfa1p has a dominant interaction over both, supports this model.

Chapter 4: Characterization of *prp43* alleles identifies three classes

4.1 Introduction to Prp43p structure

In *Saccharomyces cerev*isiae there are six essential DEAH-box helicases: the pre-mRNA splicing factors Prp2p, Prp16p, and Prp22p (Burgess *et al*. 1990; Chen and Lin 1990; Company *et al*. 1991), ribosome biogenesis factors Ecm16p and Dhr2p (Colley *et al*. 2000), and Prp43p, which functions in both processes (Martin *et al*. 2002; Lebaron *et al*. 2005; Combs *et al*. 2006; Leeds *et al*. 2006). Additionally, there is a non-essential putative DEAH-box helicase, YLR419W (Shiratori *et al*. 1999) which has yet to be characterized and has no obvious ortholog (Sanjuán and Marín 2001). To date, Prp43p is the only DEAH-box helicase for which there is a crystal structure of the full-length protein. It is composed of six domains: an N-terminal region, two RecA domains composing the RNA helicase domain, and a region common to all DEAH-box proteins (DEAH-C) (Stevens 2010) containing a degenerated winged helix domain (WHD), a ratchet domain, and a C-terminal domain (CTD) including an oligonucleotide/oligosaccharide binding (OB) fold (He *et al*. 2010; Walbott *et al*. 2010, Fig. 4.1A). While the DEAH-box helicases all share the highly conserved helicase core with other superfamily 2 (SF2) RNA and DNA helicases, the homology between them extends into their DEAH-C, indicating potential functional similarity.

The four DEAH-box helicases involved in splicing share 30% amino acid similarity in the region directly downstream of motif VI, which spans amino acids 431-720 in Prp43p (Tanaka and Schwer 2006). In fact, the C-terminal domain of hPrp22p (aa 950-1183) adopts a domain structure similar to Prp43p, including a degenerated winged-

helix domain, ratchet domain, and OB-fold. Comparison of the CTDs using the crystal structures of hPrp22 and ScPrp43, and homology modeling of hPrp16 shows a putative common RNA binding surface in the helicases (Kudlinzki *et al*. 2012). Divergent sequences or structures in the C-terminal domains of the spliceosomal DEAH-box helicases have been hypothesized to be necessary for their specific interactions and functions (Kudlinzki *et al*. 2012; Tanaka and Schwer 2006).

Like other RNA helicases, Prp43p has poor inherent unwinding and ATPase activities, as well as lacking substrate specificity (Tanaka and Schwer 2006; Walbott *et al*. 2010). Dynamic movements of its RecA domains with respect to the WHD are thought to be necessary in the course of its ATPase cycle, and may be enhanced by interactions with cofactors (Tanaka *et al*. 2007; Walbott *et al*. 2010). Prp43p can hydrolyze all NTPs and dNTPs (Tanaka and Schwer 2006), which is similar to DEAH-box helicases Prp2p (Kim *et al*. 1992), Prp16p (Schwer and Guthrie 1992), and Prp22p (Tanaka and Schwer 2005). The NTP binding site of Prp43p is formed between its two RecA domains, and RecA-1 interfaces intimately with the WHD. The Prp43p co-crystal structure with ADP indicates the adenine base stacks between Phe359 and Arg159, and though Ser155 and Gln354 coordinate water molecules to form hydrogen bonds with the base, there is no specificity for adenine (He *et al*. 2010; Walbott *et al*. 2010). By modeling Prp43p/RNA interaction using the structurally similar Ski2-like archeal DNA helicase Hel308 in complex with DNA (Büttner *et al*. 2007), the RecA-1 domain of Prp43p is predicted to bind to the 3’ end of ssRNA and the RecA-2 domain binds to the 5’ end (Walbott *et al*. 2010). A long β hairpin (5' HP, aa 396-415) existing between the

motifs V and VI of RecA-2 inserts physically between the WHD and the CTD. Putatively it blocks the binding pocket for ssRNA, so that when bound to ADP, Prp43p cannot bind RNA. The binding of ATP is predicted to induce a conformational change in Prp43p so that the 5' HP is released from the RNA binding pocket, allowing ssRNA to bind. Alternatively, when ssRNA binds to the RecA domains, a closed conformation is induced which produces an optimal ATP binding site (He *et al*. 2010), which would account for the stimulation of the ATPase activity of Prp43p by ssRNA (Tanaka and Schwer 2006). Therefore, the 5' HP is hypothesized to have two simultaneous roles: an autoinhibitory role preventing the binding of ssRNA in the ADP conformation and a role in active strand separation or RNA-protein remodeling (He *et al*. 2010). The 5' HP shows genetic interaction with the G-patch protein Ntr1p (Tanaka *et al*. 2007), and a mutation in the 5' HP has also been shown to be partially defective in 20S rRNA cleavage (Pertschy *et al*. 2009; Fig. 4.1A), indicating that this hairpin is potentially involved in the mechanism of Prp43p function in both splicing and ribosome biogenesis.

The WHD and ratchet domains (aa 455-520 and 521-634, respectively) of Prp43p have structural similarity to the Ski2-like family member Hel308 (Büttner *et al*. 2007; Walbott *et al*. 2010). Hel308 has a long α helix with two residues (Arg592 and Trp599) that have been proposed to stack on the nucleotides bases and act a hooks on the ssDNA (Büttner *et al*. 2007). Prp43p also contains equivalent arginines in this α helix (Arg611 and Arg625) that could act as the hooks for its processivity (Walbott *et al*. 2010), but their contributions to helicase activity have yet to be demonstrated. Also like Prp43p, Hel308 contains a β hairpin (5' HP) (Büttner *et al*. 2007). The proposed model is for the

ratchet α helix of Prp43p to pull ssRNA into the cavity in a 3' to 5' manner, then the β hairpin actively separates the duplex (Walbott *et al*. 2010).

In Hel308, the C-terminal domain works as an autoinhibitory domain (Richards *et al*. 2008), but in contrast this region of Prp43p has been proven to be necessary for stimulating its RNA unwinding activity and is essential *in vivo* (Walbott *et al*. 2010; Martin *et al*. 2002; Tanaka *et al*. 2007). The OB-fold in the CTD (aa 660-712) and the groove formed by the β strands has a positive electrostatic surface for interacting with nucleic acid (He *et al*. 2010; Walbott *et al*. 2010).

Though the C-terminal 35 amino acids (aa 732-767) have been shown to be essential for cell growth (Martin *et al*. 2002; Tanaka *et al*. 2007), their loss does not affect the basal ATPase activity of Prp43p. However, a mutant truncated from the OB-fold has an RNA-stimulated ATP hydrolysis rate 20-fold lower than wild-type. The decrease in RNA stimulation in the truncation is consistent with the 200-fold decrease in its RNA affinity (Walbott *et al*. 2010). Similarly, in ScPrp22p the CTD is necessary for the RNA stimulation of its helicase activity, but it also stimulates the inherent ATPase activity (Kudlinzki *et al*. 2012). Therefore, RNA binding by the OB-fold may actively participate in the ATPase cycle. Additionally, the OB-fold is may be providing some of the selectivity for the RNAs to which Prp43p binds (Walbott *et al*. 2010). The N-terminus of Prp43p wraps back towards the CTD and residues 10-36 make intimate contact (He *et al*. 2010); however, it has been shown that the N-terminal helices are not necessary for cell growth (Walbott *et al*. 2010; Martin *et al*. 2002).

The C-terminus of Prp43p also makes important contacts with a set of three G-patch proteins. Pfa1p (Walbott *et al*. 2010; see Chapter 2.1.1) or Ntr1p (Tanaka *et al*. 2007; see Chapter 3.1.1) have been shown to stimulate Prp43p to unwind helices both in the 3' to 5' or 5' to 3' direction. Additionally, Pfa1p can stimulate the ATPase activity of Prp43p, but Ntr1p does not (Walbott *et al*. 2010; Tanaka *et al*. 2007). The last 110 amino acids of Prp43p are necessary to mediate its interaction with the C-terminal domain of Pfa1p, which contains its G-patch domain. The N-terminal portion of Pfa1p can still associate with Prp43p in the absence of the C-terminus of Pfa1p, however this interaction does not stimulate the ATPase activity of Prp43p (Walbott *et al*. 2010). In contrast, the G-patch domain of Ntr1p is essential for its interaction with Prp43p. Single point mutants in the hydrophobic leucines render Ntr1p unable to bind Prp43p (Tanaka *et al*. 2007).

We sought to understand how the C-terminus of Prp43p impacts its function in either ribosome biogenesis or pre-mRNA splicing. We hypothesized that specificity domains existed downstream of the conserved helicase core to provide substrate specificity for Prp43p. Through mutagenesis of the C-terminus of Prp43p, we characterized new *prp43* mutants and determined that Gno1p and Ntr1p binding sites on Prp43p at least partially overlap.

4.2 MATERIALS AND METHODS

4.2.1 Plasmids

pRS316-*PRP43-TAP::KANMX* (pJPS1162) was isolated from yeast strain yJPS797/SS3017 (Leeds *et al*. 2006), and *PRP43-TAP* was subcloned into pRS415 using SalI and BamHI/BglII. The *PRP43* ortholog from *S. pombe* was PCR amplified from genomic DNA (Strain *SP6* from Jon Huibregtse/Susan Forsberg) using oligonucleotides S.p. Prp43 3NotI and S.p. Prp43 5BglII (Table 4.2) and cloned into pRS426 behind a *GPD* promoter. Directed *prp43* point mutants and 5' HP deletion were created by adapting Quikchange mutagenesis (Stratagene) using oligonucleotides described in Table 4.2. Plasmids containing random *prp43* alleles were isolated, transformed into *E. coli* and sequenced.

4.2.2 Construction of yeast strains

SS3105 was created by deleting *TRP1* through *MPY-URA3* replacement (Schneider *et al*. 1996). *MPY-URA3* PCR fragments were introduced via lithium acetate transformation (Gietz and Woods 2002) and transformants were selected on SD-URA. Serial dilutions in non-selective minimal media allow *MPY* repeats to recombine, resulting in the loss of *URA3*. Isolates deficient in both *TRP1* and *URA3* were then transformed with pRS316-*PRP43-TAP::KANMX*, and pRS415-*PRP43-TAP* was lost through dilution in SD-URA. 13MYC tags for each G-patch protein were incorporated into SS3105 using methods described in 2.2.1 and oligonucleotides in Table 2.2.

4.2.3 Gap repair mutagenesis

Random mutants of *prp43* were generated by degenerative PCR using 1 u of Taq polymerase (Stevens lab) with primers PRP43 GapL and PRP43 GapR (Table 4.2). pRS415-*PRP43-TAP* was digested with StuI and AscI and transformed with mutagenized *prp43* PCR fragments into a SS3076 at a 3:1 molar ratio. Gap-repaired transformants were patched on SD-LEU plates then replica plated and screened at 25˚C, 37˚C and 16˚C on SD-LEU plates containing 5'FOA. Approximately 20,000 colonies were screened, and we did not recover the same mutation more than once.

4.2.4 Serial dilutions

prp43 mutants were grown in liquid medium at 25˚C and adjusted to OD (A_{600}) of 1.0. Dilutions of 10^{-2} were made and equal volumes plated on SD-LEU or SD-LEU + 5'FOA media. Plates were photographed after 3 days of incubation at 25˚C and 37˚C, and after 6 days for growth at 25˚C and 16˚C.

4.2.5 Primer extensions

Primer extensions were conducted as described above (Chapter 2.2.7) using total RNA prepared from the time points indicated at NPT. Cultures were maintained between OD (A_{600}) 0.3-0.4 throughout the time course.

Table 4.1 Yeast strains used in Chapter 4

Strain	Relevant Gene(s)	Genotype	Source/Reference
3017/ yJPS797	pRS316-*PRP43-TAP*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3* pRS316-*PRP43-TAP:: KANMX*	(Leeds *et al*.2006)
3005	pRS415-*PRP43-TAP*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*PRP43-TAP*	This study
3010	pRS415-*prp43D638N-TAP*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*prp43D638N-TAP*	This study
3011	pRS415-*prp43L649R-TAP*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*prp43L649R-TAP*	This study
3053	pRS426-*GPD-S.pombe PRP43*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS426-*GPD-S. pombe PRP43*	This study
3105	pRS316-*PRP43-TAP trp1Δ*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ*, pRS316-*PRP43-TAP*	This study
3111	pRS316-*PRP43-TAP NTR1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, NTR1-13myc::TRP1*, pRS316-*PRP43-TAP*	This study
3112	pRS316-*PRP43-TAP PFA1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, PFA1-13myc::TRP1*, pRS316-*PRP43-TAP*	This study
3113	pRS316-*PRP43-TAP GNO1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, GNO1-13myc::TRP1*, pRS316-*PRP43-TAP*	This study
3114	pRS415-*PRP43-TAP PFA1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, PFA1-13myc::TRP1*, pRS415-*PRP43-TAP*	This study
3115	pRS415-*prp43WLI-TAP PFA1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, PFA1-13myc::TRP1*, pRS415-*prp43WLI-TAP*	This study

Strain	Relevant Gene(s)	Genotype	Source/Reference
3116	pRS415-*prp43L469R-TAP* *PFA1-myc*	*MAT* **a** *his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ*, *prp43::HIS3*, *trp1Δ*, *PFA1-13myc::TRP1*, pRS415-*prp43L469R-TAP*	This study
3117	pRS415-*prp43IRK-TAP* *PFA1-myc*	*MAT* **a** *his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ*, *prp43::HIS3*, *trp1Δ*, *PFA1-13myc::TRP1*, pRS415-*prp43IRK-TAP*	This study
3119	pRS415-*PRP43-TAP* *NTR1-myc*	*MAT* **a** *his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ*, *prp43::HIS3*, *trp1Δ*, *NTR1-13myc::TRP1*, pRS415-*PRP43-TAP*	This study
3120	pRS415-*prp43WLI-TAP* *NTR1-myc*	*MAT* **a** *his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ*, *prp43::HIS3*, *trp1Δ*, *NTR1-13myc::TRP1*, pRS415-*prp43WLI-TAP*	This study
3121	pRS415-*prp43IRK-TAP* *NTR1-myc*	*MAT* **a** *his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ*, *prp43::HIS3*, *trp1Δ*, *NTR1-13myc::TRP1*, pRS415-*prp43IRK-TAP*	This study
3122	pRS415-*prp43IRK-TAP*	*MAT* **a** *his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ*, *prp43::HIS3*, pRS415-*prp43IRK-TAP*	This study
3123	pRS415-*prp43WLI-TAP*	*MAT* **a** *his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ*, *prp43::HIS3*, pRS415-*prp43WLI-TAP*	This study
3125	pRS415-*prp43WLI-TAP* *NTR1-myc*	*MAT* **a** *his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ*, *prp43::HIS3*, *trp1Δ*, *NTR1-13myc::TRP1*, pRS415-*prp43WLI-TAP*	This study
3126	pRS415-*prp43L469R-TAP GNO1-myc*	*MAT* **a** *his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ*, *prp43::HIS3*, *trp1Δ*, *GNO1-13myc::TRP1*, pRS415-*prp43L469R-TAP*	This study
3127	pRS415-*prp43WLI-TAP* *GNO1-myc*	*MAT* **a** *his3Δ1*, *leu2Δ0*, *met15Δ0*, *ura3Δ*, *prp43::HIS3*, *trp1Δ*, *GNO1-13myc::TRP1*, pRS415-*prp43WLI-TAP*	This study

Strain	Relevant Gene(s)	Genotype	Source/Reference
3147	pRS415-*prp43G395R-TAP*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*prp43G395R-TAP*	This study
3150	pRS415-*prp43A565E-TAP* *NTR1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, NTR1-13myc::TRP1*, pRS415-*prp43A565E-TAP*	This study
3151	pRS415-*prp43A565E-TAP* *PFA1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, PFA1-13myc::TRP1*, pRS415-*prp43A565E-TAP*	This study
3152	pRS415-*prp43A565E-TAP* *GNO1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, GNO1-13myc::TRP1*, pRS415-*prp43A565E-TAP*	This study
3160	pRS415-*PRP43-TAP* *GNO1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, GNO1-13myc::TRP1*, pRS415-*PRP43-TAP*	This study
3163	pRS415-*prp43IRK-TAP* *GNO1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, GNO1-13myc::TRP1*, pRS415-*prp43IRK-TAP*	This study
3166	pRS415-*prp43A565E-TAP*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*prp43A565E-TAP*	This study
3099	*PFA1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, PFA1-13myc::URA3*	This study
3103	*NTR1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, NTR1-13myc::URA3*	This study
3107	*GNO1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, GNO1-13myc::URA3*	This study
3174	pRS415-*prp43 Q732STOP*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*prp43Q732STOP*	This study
3176	pRS415-*prp43E451D M492T-TAP*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*prp43E451D M492T-TAP*	This study

Strain	Relevant Gene(s)	Genotype	Source/Reference
3178	pRS415-*prp43G473D R493K-TAP*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*prp43G473D R493K-TAP*	This study
3179	pRS415-*prp43 F437L A502P-TAP*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*prp43F437L A502P-TAP*	This study
3186	pRS415-*prp43L469R-TAP* pRS411	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*prp43L469R-TAP* pRS411	This study
3187	pRS415-*PRP43-TAP* pRS411	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*PRP43-TAP* pRS411	This study
3190	pRS415-*prp43E451D M492T-TAP NTR1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, NTR1-13myc::TRP1*, pRS415-*prp43E451D M492T-TAP*	This study
3191	pRS415-*prp43G395R-TAP NTR1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, NTR1-13myc::TRP1*, pRS415-*prp43G395R-TAP NTR1-13myc*	This study
3192	pRS415-*prp43G473D R493K-TAP*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, NTR1-13myc::TRP1*, pRS415-*prp43G473D R493K-TAP*	This study
3194	pRS415-*prp43E451D M492T-TAP* pRS411	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*prp43E451D M492T-TAP*, pRS411	This study
3195	pRS415-*prp43G473D R493K-TAP* pRS411	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*prp43G473D R493K-TAP*, pRS411	This study
3196	pRS415-*prp43E451D M492T-TAP PFA1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, PFA1-13myc::TRP1*, pRS415-*prp43E451D M492T-TAP*	This study
3197	pRS415-*prp43G473D R493K-TAP PFA1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, PFA1-13myc::TRP1*, pRS415-*prp43G473D R493K-TAP*	This study

Strain	Relevant Gene(s)	Genotype	Source/Reference
3198	pRS415-*prp43E451D M492T GNO1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, GNO1-13myc::TRP1,* pRS415-*prp43E451D M492T*	This study
3199	pRS415-*prp43G473D R493K-TAP GNO1-myc*	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3, trp1Δ, GNO1-13myc::TRP1,* pRS415-*prp43G473D R493K-TAP*	This study
3200	pRS415-*prp43Q732STOP* pRS411	*MAT* **a** *his3Δ1, leu2Δ0, met15Δ0, ura3Δ, prp43::HIS3*, pRS415-*prp43Q732STOP*, pRS411	This study

Table 4.1 Oligonucleotides used in Chapter 4

Name	Sequence
43demut BTM	5’ AATTTGCAAAACAGTACGAATGGCAGAATCTAAATAGTCCC 3’
43demut TOP	5’ GGGACTATTTAGATTCTGCCATTCGTACTGTTTTGCAAATT 3’
5'HP TOP	5’ CGTCGTTGACCCAGGAATCTCCAAGGCTTCTGCCCAACAAAG 3’
5'HP Bottom	5’ CTTTGTTGGGCAGAAGCCTTGGAGATTCCTGGGTCAACGACG 3’
S.p.PRP43 seq-1	5’ GAAATGATTCCAACACCAGTCAGC 3’
S.p.PRP43 seq-2	5’ TGCACTGGCTTTGCTAATAGGAC 3’

Name	Sequence
S.p.PRP43 seq-3	5’ ATATCAGTTGCCAAGGTACGCTC 3’
S.p.Prp43 3Not1	5’ GCGGCCGCTTAACGAGCGTTCTTTTTTGATC 3’
S.p.Prp43 5BglII	5’ AGATCTATGGAACCAGCTCAGAAGAAACTAAGG 3’
E-C2	5’ GAAAAGGCCAGCAATTTCAAGTTAACTCCAAAGAGTATCACTCAC 3’
D-A2	5’ AAACTTACAAGCCTAGCAAGACCGCGCACTTAAGCGCAGGCCCGG 3’
scR1	5’ CGTGTCTAGCCGCGAGGAAGGATTTGTTCC 3’
U3 PE	5' CCAAGTTGGATTCAGTGGCTC 3'
U1	5' GAATGGAAACGTCAGCAAACA 3'
PRP43 GapL	5’ TACCACCGCATCAACAACAAAG 3’
PRP43 GapR	5’ CCCTGTTATCCCTAGCGGATCTGC 3’
Prp43IRK TOP	5’ GCCCTAAATACTTTGACAACGCCGCAGCGGCTCTTGCGTCTGG 3’

Name	Sequence
Prp43IRK Bottom	5’ CCCAGACGCAAGAGCCGCTGCGGCGTTGTCAAAGTATTTAGGGC 3’
Prp43WLI Top	5’ CTCGGTCAGGCCCGAAGCGGCGGCTGAAATAGCGCCTGCATAC 3’
Prp43WLI Bottom	5’ GTATGCAGGCGCTATTTCAGCCGCCGCTTCGGGCCTGACCGAG 3’
35S	5’CGCTGCTCACC AATGG3’

4.2.6 Native snRNP gels

Splicing extracts were prepared as described above from cultures shifted from 25˚C to 37˚C for 3 hours. Native gel analysis and subsequent northern blotting for the U6 snRNA was carried out as described in Chapter 2.2.9. U6 signals for U4/U6 or U4/U6•U5 were quantitated in Quantity One (BioRAD) and normalized to the wild-type signal for each species.

4.2.7 rRNA analysis

Wild-type *PRP43* and mutants were grown at 25˚C until approximately OD (A_{600}) 0.3-0.4, then shifted to 37˚C by harvesting and resuspending in pre-warmed media. Cultures were maintained at low OD (A_{600}) < 0.5 throughout the time course, and whole cell RNA was prepared from 0, 0.5, 1, 2 or 3 hours time points as described in Chapter

2.2.6. The primers used for northern blotting of the rRNAs were previously described (Combs *et al*. 2006).

4.2.8 Pulse chase analysis

Wild-type and mutant *prp43* strains harboring pRS411 were utilized for pulse chase analysis (Table 4.1). Strains SS3187, SS3186, SS3174 were grown to an OD (A_{600}) of 0.3-0.4 at 25˚C in S-MET medium supplemented with 2% dextrose. After growth in pre-warmed media for 1 hour, forty OD (A_{600}) units were harvested by centrifugation and resuspended in 12 ml of pre-warmed SD-MET medium. Pulse chase was performed as described in Chapter 2.2.7. Pulse chase analysis for SS3176 and SS3178 were begun once cultures reached OD_{600} 0.3-0.4 at 25˚C

4.2.9 TAP immunoprecipitations

For Pfa1p-myc or Gno1p-myc coimmunoprecipitations, cells were grown at 25˚C in YP media containing 2% dextrose to approximately OD (A_{600}) 0.3-0.4 before shifting to pre-warmed media and grown at 25˚C for 3 hours. For Ntr1p-myc coimmunoprecipitations, cultures were grown to OD_{600} 0.8, as for splicing extract conditions, before shifting to pre-warmed media for 3 hours. TAP purifications were carried out as described in Chapter 2.2.5.

4.2.10 RNA immunoprecipitations

Extracts were prepared as described above for 35S pre-rRNA analysis (Chapter 2.2.5), but for snRNA association, extracts were clarified with an additional centrifugation step of 100,000 x *g* for 1 hour as in splicing extract preparation (Chapter 2.2.9). In either case, TAP complexes were isolated as described above and eluted from

beads by boiling in IPP150 (10 mM Tris, pH 8.0, 150 mM NaCl, 0.1% NP-40, 1.5 mM $MgCl_2$, 8% glycerol, 1 mM DTT). Eluted material was phenol/chloroform extracted and RNAs precipitated with $1/10^{th}$ volume NaOAc under 2.5 volumes of 100% ethanol. To analyze immunoprecipitated snRNAs, 1% of RNA from total and flow-through fractions and 20% of RNA from bound fractions were separated though an 8M urea 7% acrylamide denaturing gel. Northern blot analysis was performed as previously described (Combs *et al*. 2006). For 35S RNA immunoprecipitations, primer extensions were performed as describe above (Chapter 2.2.8) with either 1% input RNA or 50% bound RNA using the 35S oligonucleotide (Table 2.2)

4.2.11 Polysome profiling

Cultures were grown at 25º C in YP + 2% dextrose media to OD (A_{600}) 0.3-0.4 before shifting to pre-warmed media and grown at 37˚C for 3 hours. OD (A_{600}) was maintained between 0.3-0.4 throughout the shift to ensure ribosome production. Cycloheximide was added at 150 μg/ mL to lock polysomes *in situ* and incubated for 10 minutes at 37º C. Cells were pelleted in ice and washed once with cold 1X PBS then once in lysis buffer (10 mM Tris-CL, pH7.5, 10 mM $MgCl_2$, 150 μg/ mL cycloheximide, 200 μg/mL heparin, 100 mM KCl, 6 mM β-mercaptoethanol). Cells were lysed by vortexing four times with 1 volume of glass beads and extracts were clarified by centrifugation at 20,000 x *g* for 10 minutes. 20 OD units (A_{260}) were layered on 7-47% sucrose gradient and centrifuged at 274,000 x *g* for 2.5 hours. Gradients were fractionated and OD (A_{254}) recorded for polysome traces.

4.3 RESULTS

4.3.1 Identification of mutations in the C-terminus of Prp43p that confer conditionality

The C-terminal 55 amino acids of Prp43p have been shown to be essential *in vivo* (Martin *et al*. 2002). This region includes the C-terminal two α helices adjacent to the OB-fold and affects the RNA-dependent stimulation of the ATPase activity of Prp43p (Walbott *et al*. 2010). We hypothesized that ATPase stimulation might occur concurrently with substrate binding and/or cofactor binding by the DEAH-C. The *in vivo* responsibilities of the DEAH-C domains of Prp43p have not been investigated. An alanine screen of amino acids 431-720 revealed 13 lethal or conditional alleles (Tanaka and Schwer 2006), although their effects in pre-mRNA splicing and ribosome biogenesis are not yet known.

We set out to identify alleles that might illuminate the *in vivo* function of the C-terminus of Prp43p using a gap-repair mutagenesis screen. The C-terminal 1.2 KB region of *PRP43* was randomly mutagenized via degenerate PCR (Fig. 4.1B) and gap-repaired in a *prp43*Δ strain harboring *Schizosaccharomyces pombe PRP43* on a *URA3* vector. By using the *S. pombe* ortholog of *PRP43* our observed frequency of gap-repair from the wild-type ectopic *S. cerevisiae PRP43* was reduced. To identify conditional *prp43* alleles, colonies were screened at 37°C and 16°C and compared to growth at 25˚C (Fig. 4.2B &C). A total of sixteen conditional alleles were isolated and sequence verified (Appendix, Table A1); however, we focused on seven containing only one or two mutations in this region (Fig. 4.2A). The only cold sensitive (cs) mutant identified,

prp43D638N, shows only a mild growth defect. Most of the mutants in our collection do not correspond to the previously identified mutants from the alanine screen (Tanaka and Schwer 2006); however, the screen did produce a truncation mutant (Q732STOP), which is very similar to one previously shown to be ts (temperature sensitive) (Martin *et al*. 2002). Also, we identified G359R, which is a mutant in the same residue as the *prp43-1* allele, G395E (Arenas and Abelson 1997), and lies eleven amino downstream of motif V at the base of the 5' HP. The *prp43G395E* allele was previously characterized as ts, though this sensitivity can be rescued by shifting back to PT (Arenas and Abelson 1997). Additionally, we directly reproduced triple alanine mutants ^{648}IRK^{650} and ^{717}WLI^{719} for further analysis (Tanaka and Schwer 2006). Some alleles, such as *prp43G395R* and *prp43L494R*, showed growth defects even at PT, which were exacerbated at NPT. Interestingly, two separate alleles *prp43-1248* and *prp43-1504* contain mutations that include two adjacent residues, M492 and R493.

Based on the structure of Prp43p, we also deleted the 5' HP to confirm its necessity *in vivo* (Fig. 4.1C, bottom panels). The deletion of this hairpin was lethal, which is consistent with the proposal that it slices through duplexes (Walbott *et al*. 2010). Though it is possible that the 5' HP also plays an autoinhibitory role to prevent the binding of ssRNA (He *et al*. 2010), its essentiality implies autoinhibition is likely not its only function.

Mapping our alleles on the known crystal structure of Prp43p (Walbott *et al*. 2010; He *et al*. 2010), we observed that we have isolated mutations in each domain (Fig. 4.3A). Unfortunately, we had not anticipated the complex structure of the RecA-2

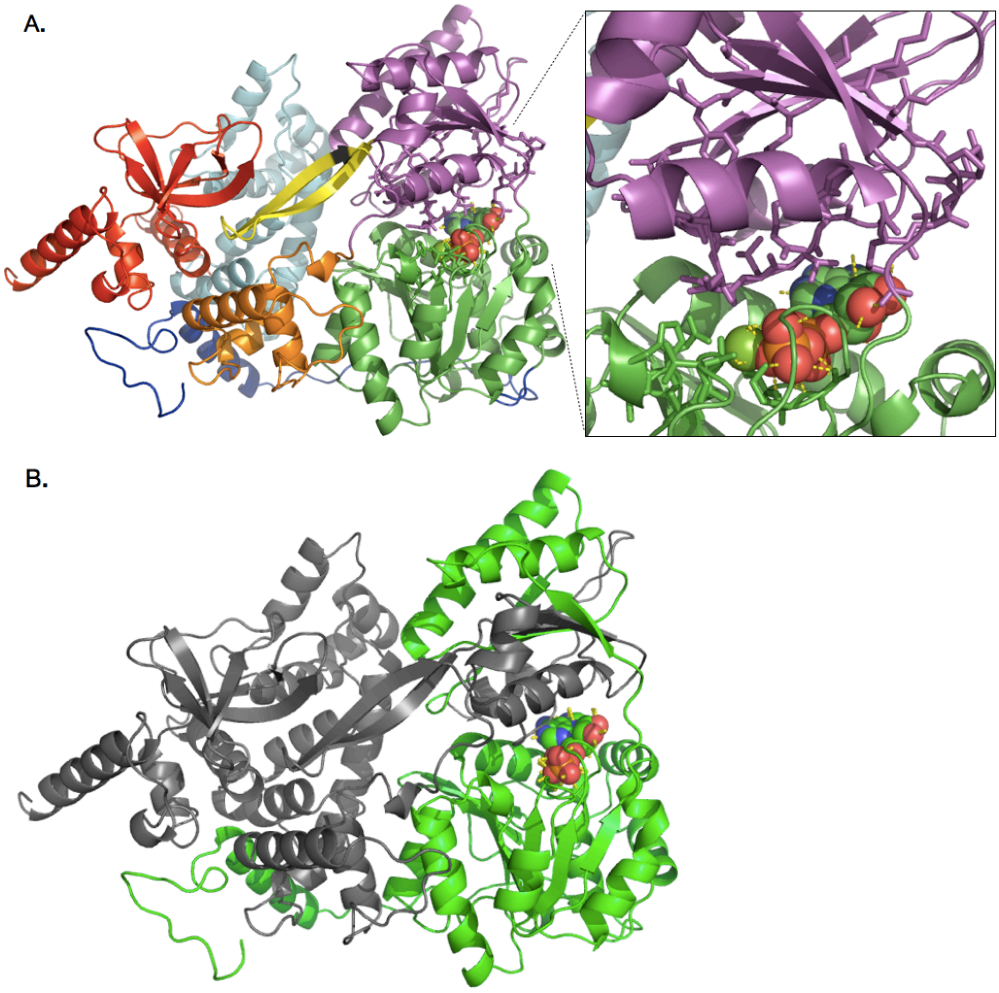


Figure 4.1 Mutagenesis of the DEAH-C of Prp43p.

A). Prp43p domains are shown as described in Walbott *et al*. and He *et al*. ADP-Mg lies between the RecA-1 (green) and RecA-2 (purple) domains. Characteristic residues of the SF2 conserved motifs are modeled as sticks, along with the polar contacts to ADP and Mg^{2+} (inset). The 5' HP (yellow) extends into the center of the enzyme. The S414F
(Figure 4.1 contd.) mutant (Pertschy *et al*. 2009; grey) maps near the base of this hairpin. The degenerate WHD (orange) lies mostly on the face opposite of the ratchet domain (cyan). At the C-terminus is the CTD (red), composing the OB-fold and ending with an extended α helix. The last five amino acids were not solved in the structure but are modeled in. The NTD (blue) contacts the CTD and the ratchet domain, and through a long unstructured linker it connects to the RecA-1 domain. File obtained from PBD (DOI:10.2210/pdb3kx2/pdb) and manipulated in MacPyMOL. B). The region of Prp43p targeted for mutagenesis (grey). Due to the complex structure of the RecA-2 domain, we unknowingly targeted a portion of that domain that could directly affect the inherent activity of Prp43p. PBD file same as in (A).

domain with the 5' HP injected between the motifs V and VI; therefore, alleles containing R370G and F437L lie closer to the conserved SF2 motifs than expected. These cluster in the upper back portion of the RecA-2 domain, and, though they are removed from the substrate binding site, it is possible that they affect the inherent activity of Prp43p.

Additionally, most of the mutations are predicted to lie on the surface of the protein (Fig. 4.3B). The *prp43IRK* and *prp43WLI* mutants line a small pocket beneath the OB-fold in the CTD, and the *prp43Q732STOP* mutant neatly truncates the α helix that also contributes to this pocket.

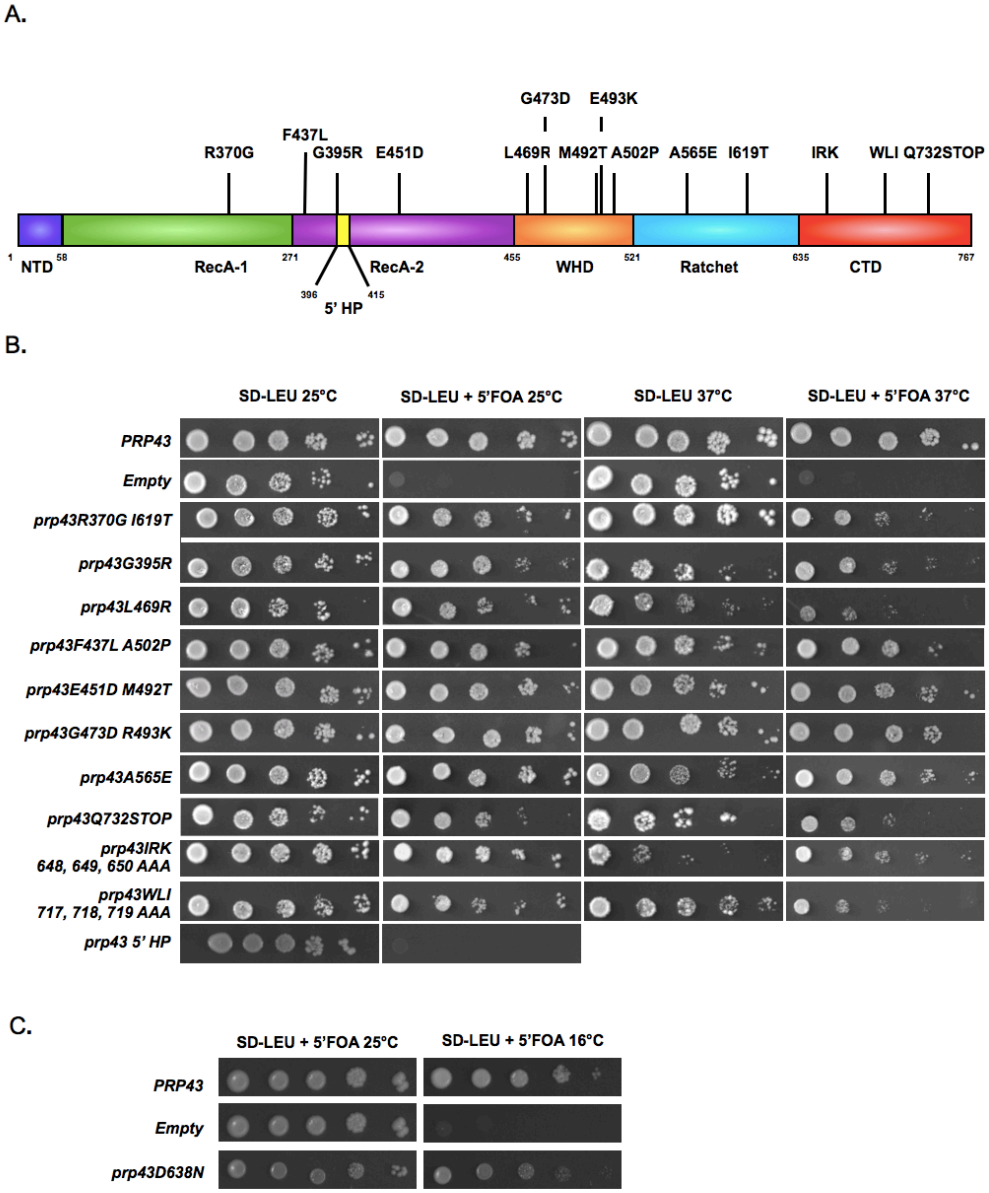


Figure 4.2 *prp43* conditional mutants isolated from the gap repair mutagenesis screen

A). Schematic diagram of mutations identified with respect to the domains of Prp43p.
B & C). Serial dilutions of ts *prp43* (C) or cs (D) alleles demonstrate growth defects at PT and/or NPT. Additionally, the *prp43 5'HP* deleting the entire 5' HP is lethal (C, bottom panels).

4.3.2 CTD mutants result in a strong first step pre-mRNA splicing defect.

To gain an understanding of the molecular basis underlying the growth phenotypes for each of the *prp43* mutants, we began by analyzing pre-mRNA splicing. Prp43p removes the lariat intron from the post-splicing lariat RNP (Martin *et al*. 2002); however, we did not identify any mutants which accumulated the U2/U6•U5 species *in vivo* (data not shown). Instead, we tested for pre-mRNA accumulation, which we have also detected for the dominant cs mutant *prp43Q423N* (Fig. 3.1B). The Q423 residue lies in motif VI, which affects ATP binding and hydrolysis. Indeed, the Q423A mutant is dominant negative, and prp43Q423Ap hydrolyzes ATP at only 13% the rate of wild-type (Martin *et al*. 2002). Notably, other alleles with mutations in the helicase core, *prp43S247A* and *prp43G429A*, which lie in motifs III and VI, respectively, do not lead to this splicing defect (Combs *et al*. 2006). Despite the lack of a splicing defect, the S247A mutation reduces ATP hydrolysis by 35% (Martin *et al*. 2002), and a mutation to leucine in the same position renders a dominant negative phenotype (Arenas and Abelson 1997).

Through primer extension analysis for the pre-snoRNA U3 we found that mutants in the CTD strongly accumulate pre-U3 snoRNA, indicating they have a defect in the first-step of pre-mRNA splicing (Fig. 4.4A). Specifically, *prp43IRK* and *prp43Q732STOP* accumulated U3a precursors up to 4- and 8- fold greater than wild-type at NPT, respectively (Fig. 4.4B). The pre-mRNA splicing defect observed in the CTD mutants was less pronounced than the *snu66Δ* control, which is expected considering Snu66p is directly involved in tri-snRNP assembly rather than indirectly, like Prp43p. However, not all *prp43* mutants, such as *prp43L469R* and *prp43A565E*, displayed this pre-mRNA

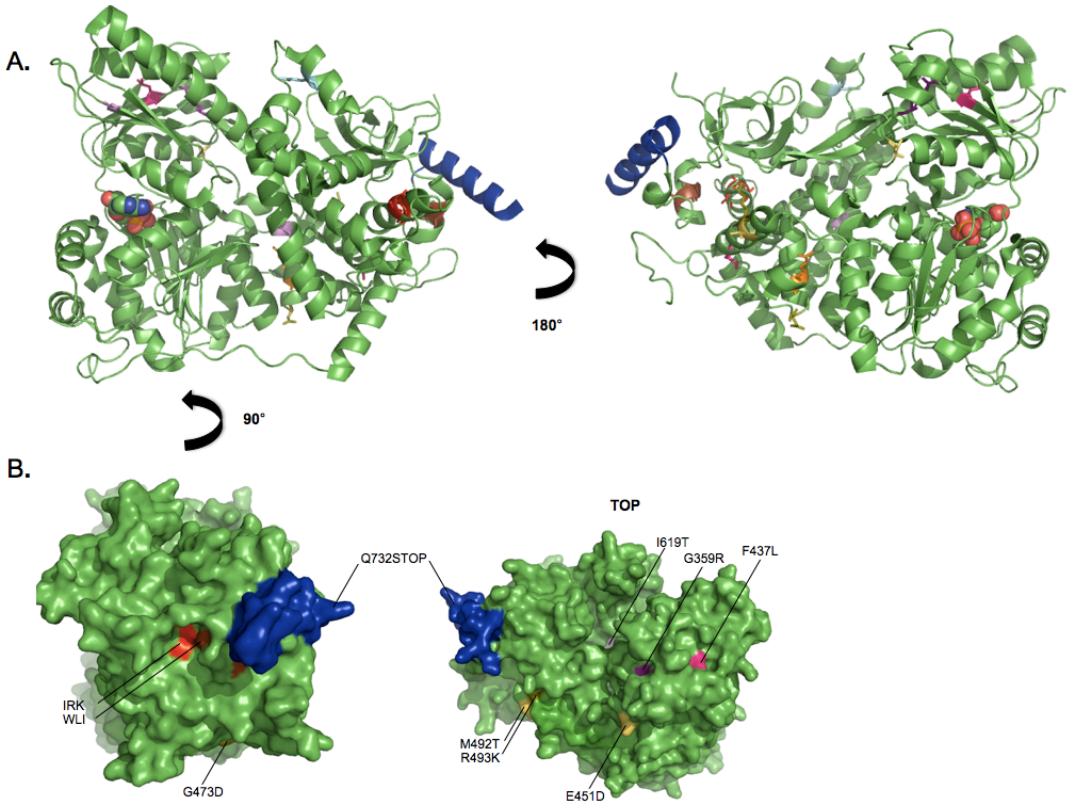


Figure 4.3 *prp43* mutants mapped on Prp43p crystal structure.

A). All *prp43* mapped on a cartoon model of the Prp43p crystal structure (He *et al*. 2010; DOI:10.2210/pdb3kx2/pdb). Each allele is colored as follows: *prp43R370G I619T*, light lavender; *prp43G395R*, purple; *prp43F467L A502P*, pink; *prp43L469R*, orange; *prp43 1248*, yellow; *prp43-1504*, olive; *prp43A565E*, light blue; *prp43IRK*, red; *prp43WLI*, burnt orange; *prp43Q732STOP* truncation, blue. B). Most *prp43* alleles map to predicted surface residues. The IRK, WLI and Q732STOP mutants cluster in the CTD and make a pocket. Along the WHD face both mutations in the alleles *prp43E451D M492T* and *prp43G473D R492K* lie on the surface. From the top, the mutations near the RecA-2 domain, F437L and G395R are visible. Other mutated residues, R370G, A565E and I619T, are predicted to be on the surface, while L469R and A502P are not (not depicted).

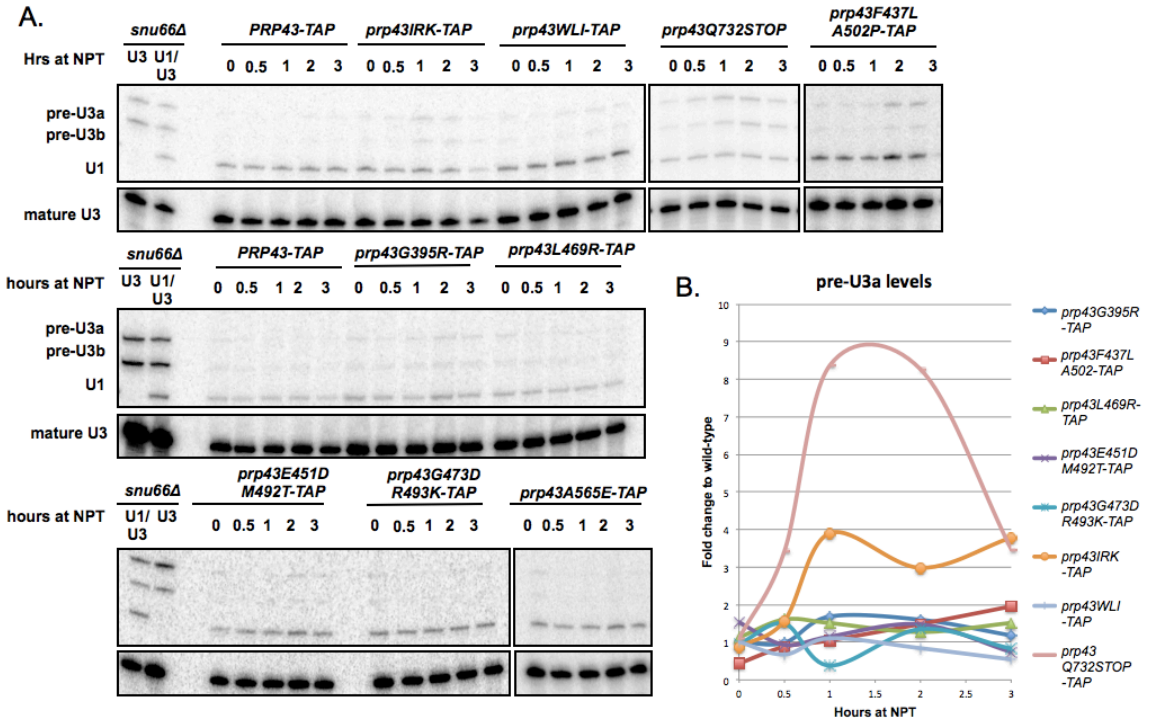


Figure 4.4 pre-U3 snoRNAs accumulate in *prp43IRK* and *prp43Q732STOP* mutants.
A). Primer extension analysis for the pre-U3 snoRNAs in each of the *prp43* mutants after shift to NPT. RNAs from each time point were primed with U3 or U1 oligonucleotides. Both U3a and b precursors and the mature mRNA are visualized using the same U3 probed. Extensions for the U1 snRNA are utilized as an internal loading control.
B). Quantitation of audioradiographs in (A). pre-U3a signals were normalized to the U1 internal control before computing fold change compared to the shifted wild-type control. Similar results are observed when U3a signals are normalized to time 0 for each mutant (not shown).

accumulation phenotype, making them more similar to the *prp43S247A* or *prp43G429A* mutants. Moreover, our *prp43G395R* allele did not drastically increase pre-U3 snoRNA levels, which might have been predicted based on the results for *prp43G395E*.

4.3.3 CTD mutants have imbalances in snRNP levels.

The loss of ATPase function in *prp43Q423N* results in snRNP imbalance so that U5 snRNPs levels are decreased resulting in a corresponding increase in U4/U6 di-snRNP levels (Fig. 3.3). Proper snRNP recycling is also dependent on the association of Prp43p with Ntr1p (Fig. 2.12 &13; Fig. 3.3; Boon *et al*. 2006). We hypothesized that mutants that accumulated pre-U3 snoRNAs would also be deficient in tri-snRNP, as in *prp43Q423N* at NPT. We further analyzed the snRNP composition in each mutant after growth at NPT for 3 hours. snRNPs were analyzed on a native gel by Northern blotting for U6 snRNA. Since the U4/U6 di-snRNP is present in low concentrations under wild-type conditions, slight changes in the abundance of this species is easier to identify in the mutants than differences in the highly abundant tri-snRNP. Indeed, we observed that the CTD alleles (*prp43IRK*, *prp43WLI* and *prp43Q732STOP*) have a strong increase in the steady-state U4/U6 snRNP levels with slight measurable decreases in tri-snRNP (Fig. 4.5). Our data are supported by previous work that shows the WLI surface is important for Prp43p binding to Ntr1p (Tanaka *et al*. 2007). In contrast, *prp43* alleles that did not show first step pre-mRNA splicing defects, *prp43L469R*, *prp43G395R*, and *prp43A565E*, had only minor effects on U4/U6 snRNP levels. Once again, this is in contrast to what might be expected for *prp43G395R* considering the previously mentioned data for *prp43G359E*. Splicing activity assays showed that *prp43G395E* results in increased pre-

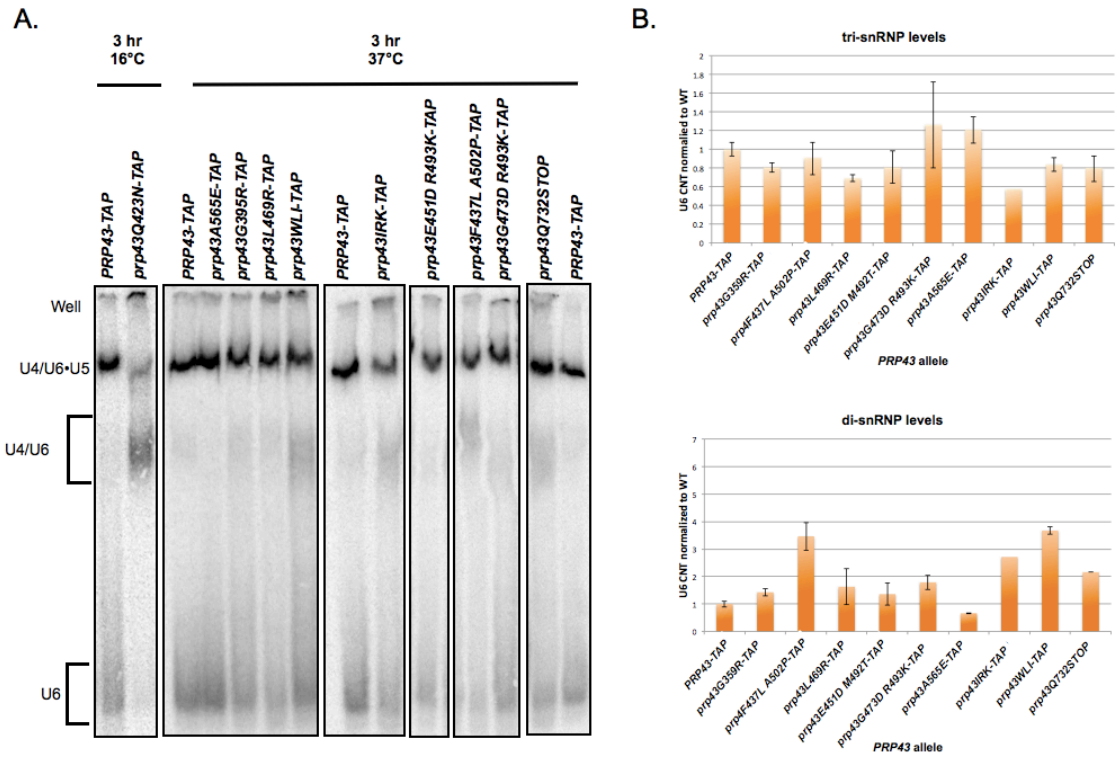


Figure 4.5 *prp43* mutants with pre-U3 accumulation also have defects in snRNP recycling.

A). Representative native gels probed for U6 snRNA to visualize U4/U6•U5, U4/U6, and mono-U6 species in *prp43* mutants at NPT. snRNP species are indicated to the left, and multiple samples of the wild-type are included as a reference. The ATPase allele prp43 Q423N is included for comparison.
B). Quantitation of the U6 counts for either U4/U6 snRNP or tri-snRNP species were determined in Quantity one and normalized to wild-type.

mRNA levels due to a failure to dissociate the spliceosome (Arenas and Abelson 1997). While Prp43p has also been shown to be able to trigger the discard of splicing intermediates at either the first or second step of pre-mRNA splicing (Koodathingal *et al*. 2010), under wild-type conditions we do not think this role would result in a defect leading to conditionality.

4.3.4 *prp43* mutants range from milder to more severe ribosome biogenesis defects than ATPase mutants.

In addition to the ATPase mutants *prp43S248A*, *prp43G429A* (Combs *et al*. 2006), and *prp43Q423N* (Leeds *et al*. 2006) that were characterized as deficient in ribosome biogenesis, one other *prp43* allele has been linked to ribosome biogenesis, *prp43S414F* (Pertschy *et al*. 2009). S414 lies in the 5' HP (Fig. 4.1A) and is hypothesized to be partially responsible for RNA unwinding (He *et al*. 2010). However, the effects of this mutation on rRNA processing are only seen with the additional loss of Ltv1p, a 20S biogenesis factor and proposed small subunit export adaptor (Pertschy *et al*. 2009; Seiser *et al*. 2006). The 20S pre-rRNA accumulates in the cytoplasm in the *ltv1Δ* mutant, and the accumulation is exacerbated in the *ltv1Δ prp43S414F* double mutant (Pertschy *et al*. 2009). In contrast, the *prp43S248A*, *prp43G429A* and *prp43Q423N* mutants are deficient in early cleavages of the 35S pre-rRNA, resulting in its transient accumulation and drastic turnover so that both 20S and 27S pre-rRNAs are depleted (Leeds *et al*. 2006; Combs *et al*. 2006). Previously, *prp43IRK* and *prp43WLI* were shown to relieve the dominant negative effect of *prp43T123A*, and, therefore, hypothesized to mediate the interaction of Prp43p with the spliceosome (Tanaka *et al*. 2007). However,

the effects of *prp43IRK* and *prp43WLI* on ribosome biogenesis have not been investigated, and considering the essential role of Prp43p in ribosome biogenesis, perhaps this pathway is partially involved in their suppression of *prp43T123A*.

For each *prp43* allele we analyzed rRNA processing by Northern blotting using D-A2 and E-C2 probes to visualize each species (Fig. 4.6; Combs *et al*. 2006). We observed four classes of ribosome biogenesis defects. The most severe effect was observed in *prp43L469R*, in which the steady-state levels of both the 20S and 27S precursors were reduced to ~25% that of wild-type. Interestingly, there was no corresponding increase in the 35S precursor, making ribosome biogenesis defects in *prp43L469R* distinct from the *prp43S247A* or *prp43G429D* mutants (Combs *et al*. 2006). In contrast, *prp43IRK,* and to an extent *prp43WLI*, phenocopied the ATPase mutants, with strong increases in the 35S precursor and decreases in the downstream products. Unlike *prp43L469R*, we noted the presence of a 23S aberrant species in these mutants, as well as *prp43Q732STOP*. Considering the stronger effect on pre-mRNA splicing in these CTD mutants, this ribosome biogenesis defect could be partially due to the incomplete splicing of U3. However, it is peculiar that *prp43Q732STOP* should have such a profound increase in 35S pre-rRNA levels, but not a corresponding decrease in the downstream products, as seen in *prp43IRK* and *prp43WLI*. Additionally, the double mutants in the WHD, *prp43E451K M492T* and *prp43G473D R493K*, had increases in the steady-state levels of the 35S precursor, but this did not translate to downstream decreases in the 20S or 27S pre-rRNAs. These two alleles showed a mild phenotype that

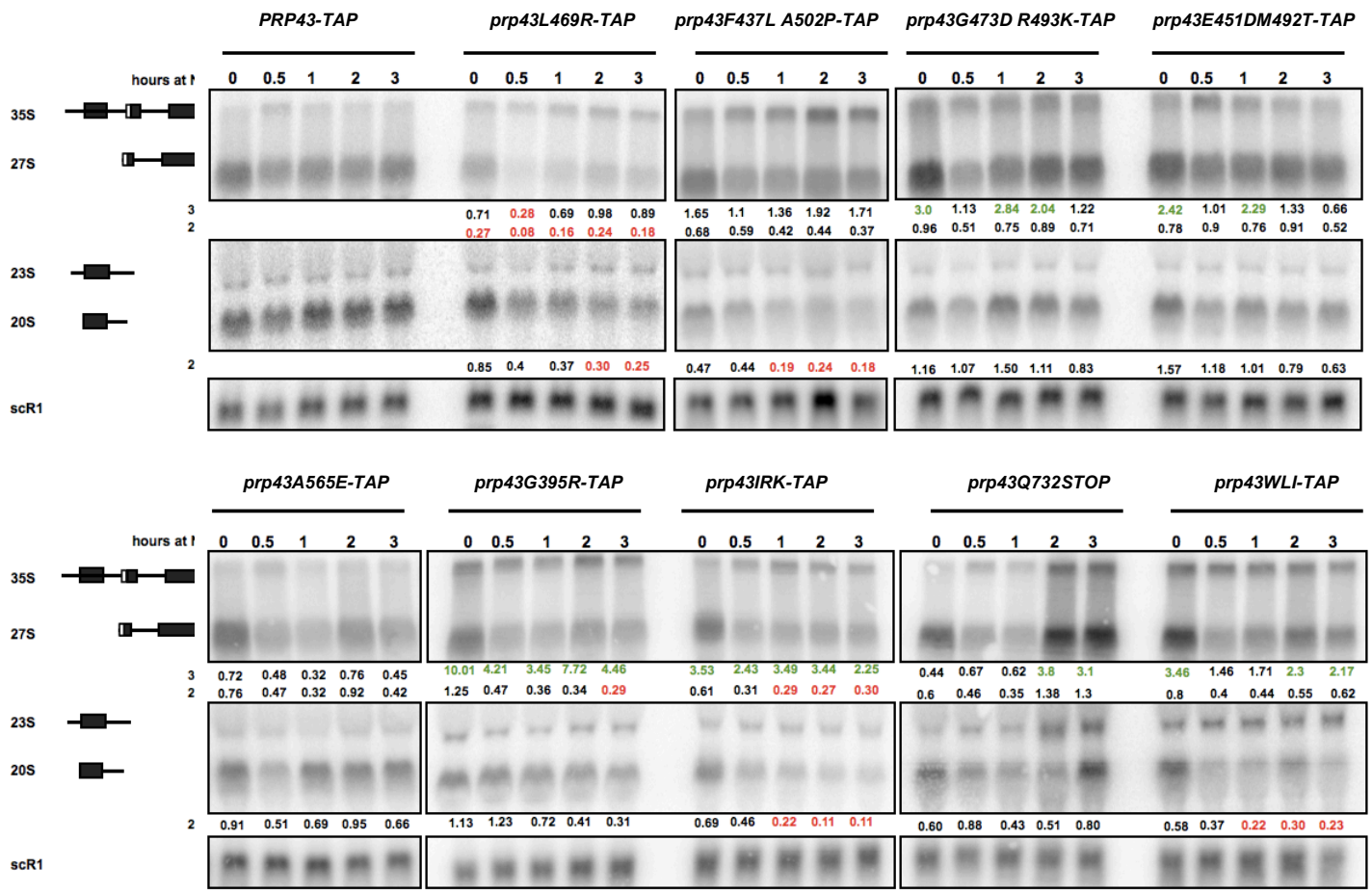


Figure 4.6 Mutations in CTD or WHD have different rRNA processing defects.

(Fig. 4.6 contd.)

The levels of steady-state rRNAs are perturbed in *prp43* mutants. RNAs were harvested at the indicated time point after shift to 37° C, and pre-RNAs were detected by northern blotting using either D-A2 or E-C2 probes (Combs *et al*.2006). A probe for scR1 is used as a loading control. For quantitation, all signals were normalized to the scR1 loading control before determining fold change compared to wild-type at the indicated time point. By comparing mutant steady-state rRNA levels to the corresponding wild-type time point we account for the effect of heat shock on rRNA processing, which is evident at 30 min in the wild-type. Quantitative increases of over 2.0 are colored green, and decreases under 0.3 are colored red.

is distinct from the CTD mutants or ATPase mutants. Finally, *prp43A565E* did not have any significant effect on ribosome biogenesis.

4.3.5 WHD *prp43* mutants have slowed rRNA processing kinetics.

Because the effect of some *prp43* mutants on ribosome biogenesis was ambiguous by Northern blot, we further investigated using pulse chase analysis. First, for *prp43L469R* we would expect to see an accumulation of the 35S rRNA precursor considering the loss of the downstream pre-rRNAs. By pulsing with radiolabeled [^{3}H]methyl-methionine and chasing with cold methionine, we detected a slight 35S pre-RNA signal that did not persist (Fig. 4.7A). Only trace amounts of downstream products were seen, even at 20 minutes. Therefore, a L469R mutation leads to more severe rRNA defects than either S247A or G429A. The 35S precursor was quickly turned over in the *prp43L469R* allele.

While we were not able to capture the reductions in 20S and 27S pre-rRNAs by northern blot in *prp43Q732STOP*, by pulse chase we observed their processing was reduced (Fig. 4.7B). Even at 20 minutes, very little 20S and 27S precursors or 18S and 25S mature rRNAs were evident. The 35S accumulation we observed by northern blotting was evident by pulse chase. Together, the pre-mRNA splicing and ribosome biogenesis defects in *prp43Q732STOP* most closely resemble that of *prp43Q423N* (see Fig. 3.1B; Leeds *et al*. 2006).

We also further questioned how *prp43E451K M492T* and *prp43G473D R493K* accumulates 35S but no changes were evident in the downstream products. We noted that even at permissive temperature, the WHD double mutants both accumulated 35S pre-

rRNA. Pulse chase analysis in each of these mutants revealed that the 35S pre-rRNA persisted, partially processed into 27S and 20S pre-rRNAs, and further matured into 25S and 18S rRNAs, respectively (Fig. 4.7C). Overall, the final levels of the mature rRNAs were reduced in the WHD mutants. The rate of ribosome biogenesis is generally slowed in each of these mutants, and this partial effect on ribosome biogenesis is distinct from any previously characterized *prp43* mutant.

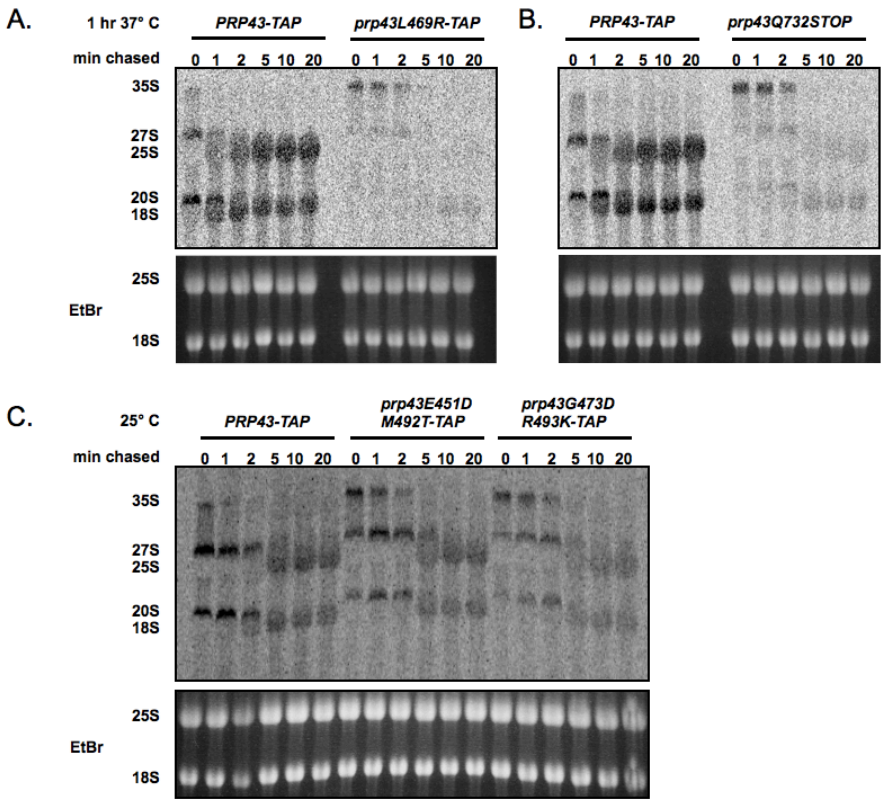


Figure 4.7 Pulse chase analysis for *prp43* mutants.

Audioradiographs are shown for tritiated RNAs isolated from the time points indicated during cold methionine chase in *prp43L469R* (A) and *prp43Q732STOP* (B) mutants after growth at 37° C for 1 hour, or in WHD mutants (C) at 25° C.

4.3.6 *prp43* mutants outside of the ATPase domain have aberrant polysome profiles

The rRNA processing defects of *prp43* ATPase mutants result in aberrant levels of mature ribosomal subunits and polysomes. In *prp43G429A* and *prp43S247A*, characteristic halfmers are present, indicating the loss of the 60S subunits (Combs *et al*. 2006). We assayed our mutants possessing variable rRNA processing defects to determine their effect on mature subunits and translation (Fig. 4.8). In the *prp43L469R* mutant (Fig. 4.8B), which has the most severe ribosome biogenesis defect, we observed decreased 80S and polysome peaks, and, similar results were seen in the *prp43* truncation mutant (Fig. 4.8C). Though the ratio of 40:60S subunits in both mutants were not observably different from wild-type, halfmers were present in the 80S and polysome peaks. Both L469R and Q732STOP mutants obviously contain aberrant subunits that result in decreased subunit joining and loading on mRNAs. In contrast, the WHD double mutants did not show any obvious changes in 80S or polysomes (Figs. 4.8D & E). The overall levels 80S and polysomes were slightly reduced, in agreement with the pulse chase analysis of these mutants (Fig. 4.8D & E). These WHD mutants do not appear to have substantial defects in translation.

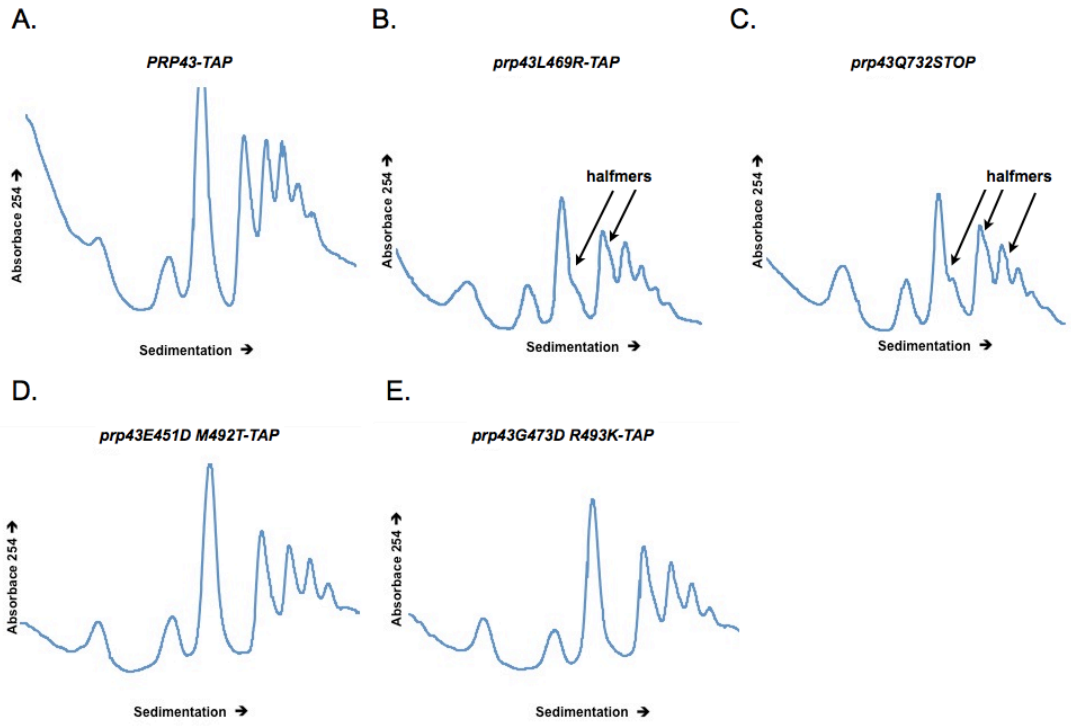


Figure 4.8 *prp43* mutants have less severe defects in mature ribosomal subunits and polysomes.

Polysome profiles in *PRP43-TAP* wild-type (A), *prp43L469R-TAP* (B), *prp43Q732STOP* (C), *prp43E451D M492T-TAP* (D), or *prp43G473 R493K-TAP* (E) after shift to NPT. *prp43 L469R* and *prp43Q732STOP* mutants both show aberrant 80S and polysome peaks, while double WHD mutants do not. Peaks from left to right: 40S, 60S, 80S, polysomes.

4.3.7 Double WHD mutants accumulate snR41 on pre-ribosomal particles.

Prp43p plays a role in removing certain snoRNAs either directly from pre-ribosomal particles or indirectly by driving the rearrangements necessary for their removal (Bohnsack *et al*. 2009). If the ATPase activity of Prp43p is defective, four snoRNAs are retained in pre-ribosomal complexes: snR39, snR39b, snR50 and snR59, all of which have binding sites on the large subunit (Bohnsack *et al*. 2009; Fig. 1.2). In *prp43* ATPase mutants, the distributions of the U14 and snR41 snoRNAs shift from mostly 40S and 60S fractions to 80S/90S and polysome fractions (Bohnsack *et al*. 2009; Fig. 1.1). Due to the mild phenotype observed in *prp43E451K M492T* and *G473D R493K*, we sought to determine if snoRNAs are not properly removed in these alleles. We separated ribosomal subunits through sucrose gradients then northern blotted for snR41, snR39, and snR50. We observed that snR41 consistently accumulated in higher sucrose fractions in the mutants compared to wild-type (Fig. 4.9). In contrast, we were unable to consistently observe the same change in sedimentation in snR39 or snR50 (Fig. 4.9 and data not shown). To account for any decrease in the total subunit levels in the mutants, we utilized 5.8S rRNA as a loading control. Both WHD mutants have a unique mild rRNA processing defect and they are unable to remove snR41 from pre-ribosomes, which is distinct from the *prp43* ATPase mutants (Bohnsack *et al*. 2009).

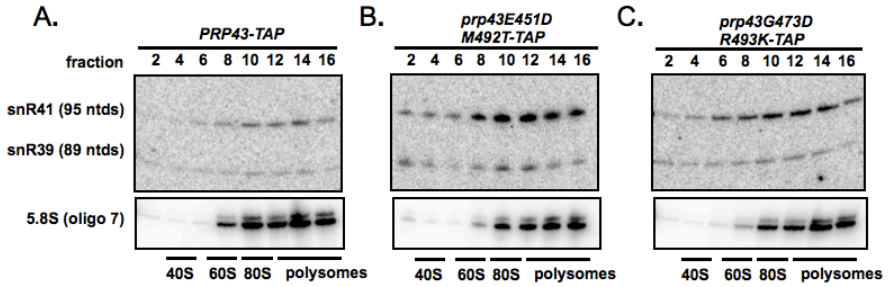


Figure 4.9 WHD mutants do not properly remove snR41.

RNA fractions from *PRP43-TAP* wild-type (A), *prp43E451D M492T-TAP* (B) or *prp43G473D R493K-TAP* (C) polysome profiles were northern blotted for snR41 or snR30 (top panels). For RNA loading control, 5.8S was also visualized by Northern blotting (bottom panels). Fractions encompassing pre-ribosomal subunits or polysome peaks are indicated below the northern blots.

4.3.8 *prp43* mutants still associate with Pfa1p-myc.

The G-patch protein Pfa1p makes extensive contacts with Prp43p and can stimulate both its ATPase and helicase activities (Walbott *et al*. 2010). We hypothesized that *prp43* mutants lose or gain association with Pfa1p at NPT, which might contribute to the observed molecular phenotypes in each allele. An increase of Pfa1p association would be consistent with Pfa1p overexpression phenotypes (see Chapter 2) but would be unlikely due to the temperature sensitive nature of the mutants. Additionally, we would not anticipate that the loss of Pfa1p association would have a drastic effect on cell growth considering that Pfa1p is non-essential, and we would not expect its robust contacts with Prp43p to be disrupted by individual amino acid changes. Indeed, none of the *prp43* mutants showed a decrease or obvious increase in Pfa1p-myc association at NPT (Fig. 4.10). Therefore, it is unlikely that the RNA processing defects observed in each mutant are linked to Pfa1p.

4.3.9 *prp43* mutants do not associate with Gno1p at NPT.

Another G-patch protein, Gno1p, has also been reported to stimulate the activity of Prp43p (Y. Henry, unpublished results) and sediments with Prp43p in 90S particles (Lebaron *et al*. 2009). We hypothesized that *prp43* mutants with defects in ribosome biogenesis might have an aberrant relationship with Gno1p. We performed immunoprecipitations with TAP-tagged constructs of each mutant after growth at NPT for 3 hours and western blotted for the presence of Gno1p-myc. Interestingly, Gno1p-myc association was lost at NPT in the *prp43G395R*, *prp43L469R*, *prp43IRK*, and *prp43WLI* alleles, but the double WHD mutants still associated with Gno1p (Fig. 4.11A).

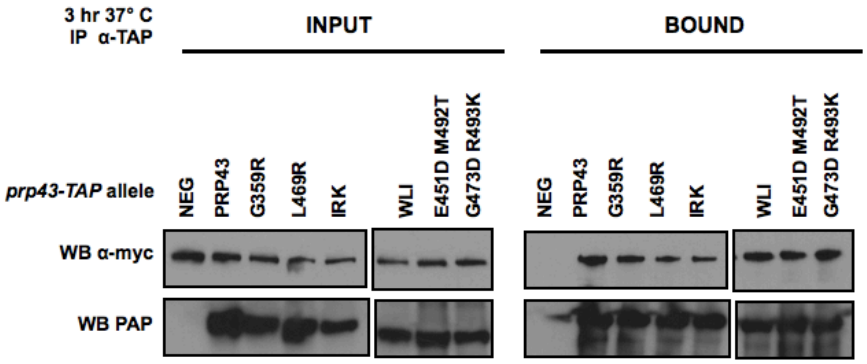


Figure 4.10 *prp43* mutants do not lose association with Pfa1p-myc.

After growth at NPT, each *prp43* mutant still immunoprecipitates with Pfa1p-myc. Complexes were precipitated using TAP tag on *PRP43* wild-type or mutants, and serially western blotted for Pfa1-myc (top panels) or Prp43p-TAP (bottom panels). Pfa1-myc extract was used as negative control for anti-TAP immunoprecipitation.
1% input; 20% bound

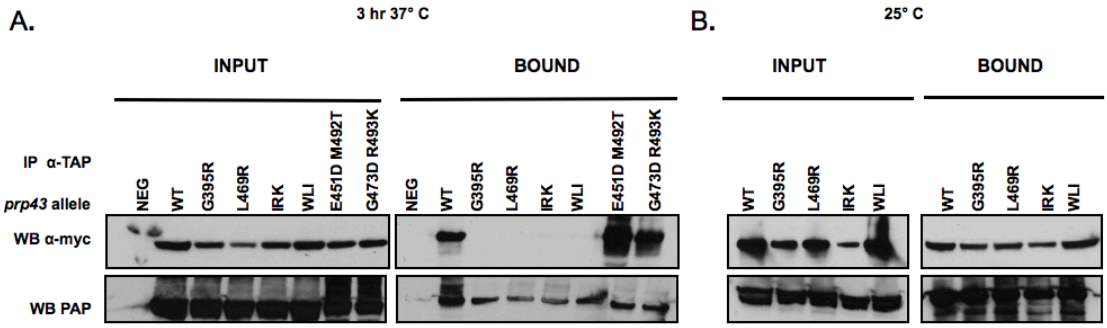


Figure 4.11 *prp43* mutants do not bind Gno1p-myc at NPT.

A). After growth at NPT for 3 hours, PRP43-TAP or *prp43*-TAP complexes were immunoprecipitated and analyzed for Gno1p-myc association by serial western blot (myc, top panels; TAP, bottom panels). Gno1-myc extract was utilized as negative control. B). Prp43 alleles that showed loss of Gno1p association at 37° C were analyzed for Gno1p association at 25° C.
1% input; 20% bound

For each mutant that was affected at 37°C, we verified that Gno1p binding was not constitutively abrogated. Indeed, at 25° C all four *prp43* mutants still associated with Gno1p-myc (Fig. 4.11B). These data indicate that the loss of Prp43p/Gno1p association partially drives the growth and ribosome biogenesis defects observed in these mutants.

4.3.10 *prp43* mutants constitutively lose association with Ntr1p

Ntr1p, stimulates the helicase activity of Prp43p (Tanaka *et al*. 2006) and is necessary for the recruitment of Prp43p to the spliceosome (Tsai *et al*. 2005). We hypothesized that *prp43* mutants that are defective in pre-mRNA splicing are unable to properly associate with the spliceosome due to a loss of interaction with Ntr1p. After growth at NPT for 3 hours, immunoprecipitations of each *prp43* allele were performed, then Ntr1p-myc binding was assayed by western blotting. Interestingly, the *prp43* alleles that lost Gno1p binding at NPT (Figure 4.12) also did not associate with Ntr1p at NPT. However, when we assayed Ntr1p binding in these mutants at 25˚C, we observed that Ntr1p association was also strongly reduced at PT. Therefore, *prp43G395R*, *prp43L469R*, *prp43IRK*, and *prp43WLI* mutants all have a constitutive defect in forming Prp43p/Ntr1p complexes. Though it might be expected that cells would not tolerate the loss of Ntr1p/Prp43p complexes, cells that lose Ntr1p/Prp43p association due to a L68A mutation in the G-patch domain of Ntr1p exhibit no growth defect at 30˚C. Only at 37˚C is the temperature sensitivity of the *ntr1L68A* allele revealed (Tanaka *et al*. 2007). In contrast, the WHD mutant *prp43G473D E493K* and the ratchet mutant *prp43A565E* still associated with Ntr1p.

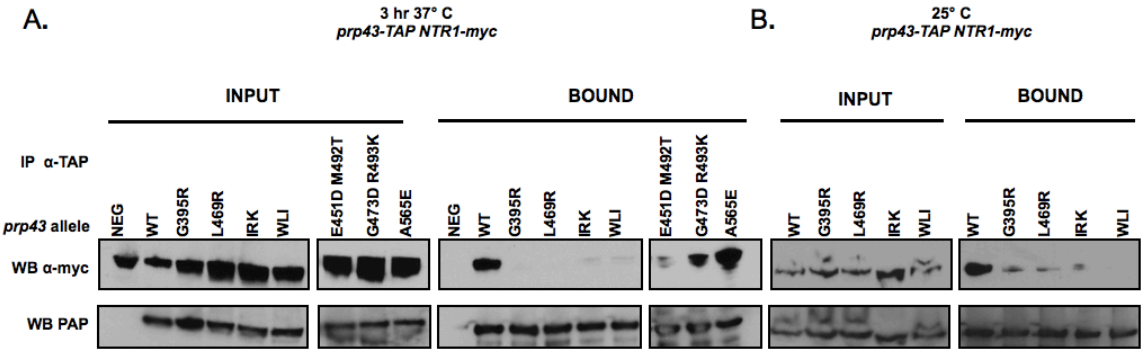


Figure 4.12 *prp43* mutants constitutively do not bind Ntr1p-myc.

A). After growth at NPT for 3 hours, PRP43-TAP or *prp43*-TAP complexes were immunoprecipitated and analyzed for Ntr1-myc association by serial western blot (myc, top panels; TAP, bottom panels). Pfa1p-myc extract was utilized as negative control.
B). Even at PT, prp43p-TAP mutants do not associate with Ntr1p-myc.
1% input; 20% bound

4.3.11 *prp43* mutants that do not associate with Gno1p also do not associate with 35S pre-rRNA

Prp43p was previously shown to retain association with the 35S pre-rRNA in the absence of Gno1p; however, the extracts used for these immunoprecipitations were also deficient in Pfa1p (Lebaron *et al*. 2009). Since our results indicated that Pfa1p could inhibit Prp43p recruitment to the 35S pre-rRNA (Fig. 2.8), we reasoned Gno1p could still be partially responsible for bringing Prp43p into 90S pre-ribosomal particles. Since PRP43p/Gno1p interaction was affected in multiple *prp43* alleles (Fig. 4.11), we analyzed *prp43* mutant association with the 35S pre-rRNA. After shift to NPT for 3 hours, the wild-type or mutant Prp43p-TAP was immunoprecipitated and the presence of the 35S pre-rRNA was determined by primer extension analysis (Fig. 4.13A). At NPT, 35S association was reduced in *prp43G395R*, *prp43L469R*, *prp43IRK* and *prp43WLI*. In contrast, 35S association was unchanged in *prp43E451D M492*T, *prp43G473D R493K*, and *prp43A565E*. The relative levels of 35S pre-rRNA bound was normalized to the respective input and compared to wild-type (Fig. 4.13B). We note that the same *prp43* mutants that lost association with Gno1p also lost association with the 35S pre-rRNA. Since the CTD of Prp43p near the OB-fold has been shown to be necessary for RNA binding (Walbott *et al*. 2010), the triple alanine IRK and WLI mutations might disrupt the RNA binding capability of this domain. Additionally, neither of the WHD mutants lost association with the 35S pre-rRNA, which is consistent with their retention of Gno1p binding (Fig. 4.11).

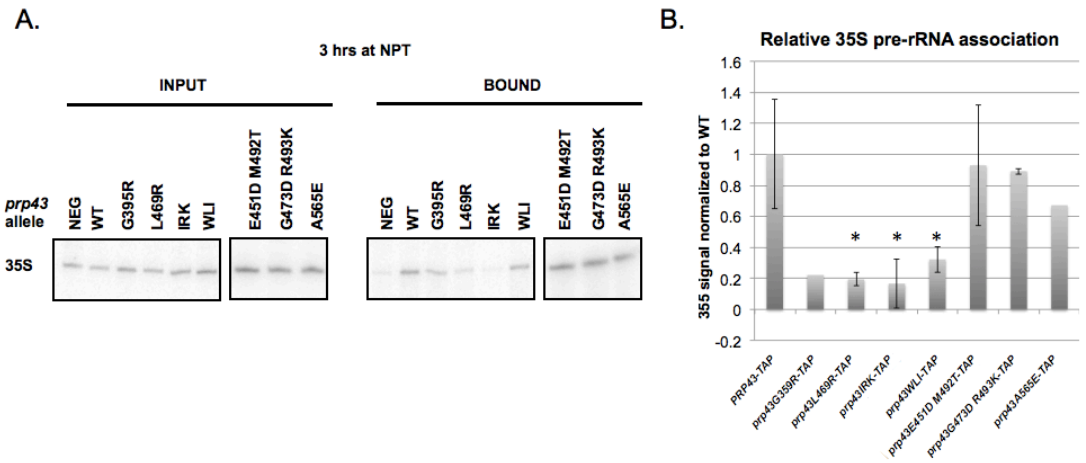


Figure 4.13 *prp43* mutants impact association with the 35S pre-rRNA.

A). Primer extension analysis for the 35S pre-rRNA immunoprecipitated with each prp43p-TAP mutant indicated. 1% input; 50% bound; NEG, untagged control.
B). Quantitation of averaged 35S pre-rRNA immunoprecipitated by each prp43p-TAP mutant. 35S pre-rRNA signals were normalized to input then fold change from wild-type determined. Dashed line represents wild-type level. Asterisks (*) indicate samples with significant p-values (p=<0.021) as determined by a two-tailed t-test.

4.3.12 *prp43* mutants G359R and L469R mutants do not interact with snRNAs

In addition to the role of the CTD of Prp43p in binding RNA (Walbott *et al*. 2010), Ntr1p has been implicated in interacting with this region (Tanaka *et al*. 2007). We would expect that in the absence of Ntr1p binding, as we observed with the IRK and WLI mutants (Fig. 4.12), Prp43p would not be able to join the spliceosome and associate with snRNAs. Indeed, these two mutations reduced the association of all snRNAs with Prp43p (Fig. 4.14A, lanes 15 &18). Additionally, Ntr1p has been shown to bind a region that encompasses a portion of the RecA2 domain and almost all of the β-hairpin (He *et al* 2010; Walbott *et al*. 2010), specifically between amino acids 390 and 411 (Tanaka *et al*. 2007). We found that the G395R mutation also abrogates snRNA association (Fig. 4.14A, lane12), which is consistent with the loss of Ntr1p association. Likewise, we found that the L469R mutation reduced Prp43p association with the snRNAs within one hour after shift to NPT (Fig. 4.14B, lane 18). Since the L469R mutation affected Ntr1p binding at PT, we also analyzed the ability of the mutant to bind snRNA at PT. As expected, the L469R mutation resulted in the constitutive inability to bind snRNAs (Fig. 4.14B, lane 15). In contrast to these mutants, prp43A565Ep, which still associates with Ntr1p-myc (Fig. 4.12B), still interacts with the snRNAs as expected (Fig. 4.14A, lane 9). We note that the association of the U5 snRNA in each mutant is the most affected compared to U2 or U6.

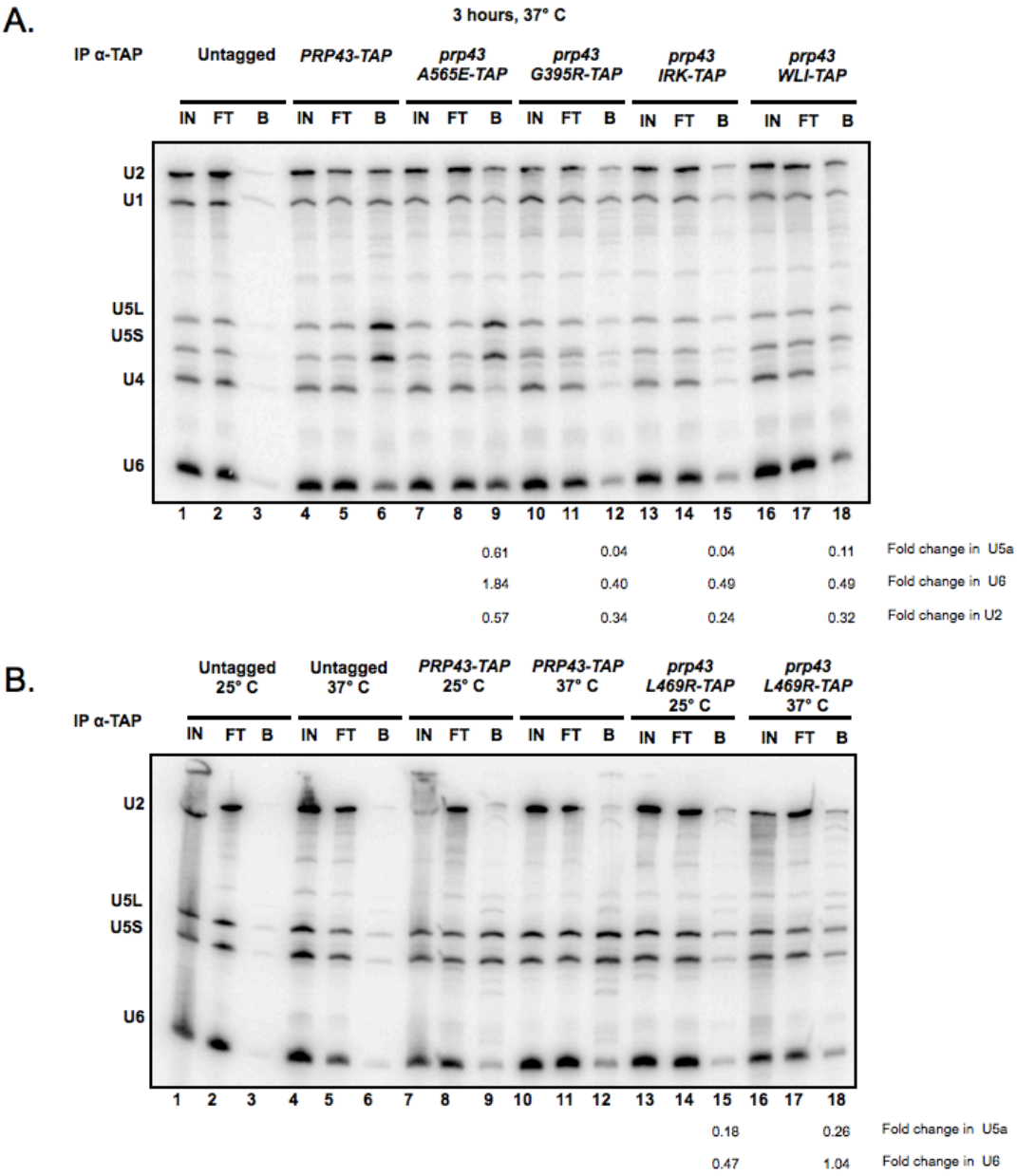


Figure 4.14 *prp43* mutants do not associate with snRNAs.

A). Northern blot analysis for snRNAs immunoprecipitated with either wild-type Prp43p-TAP or mutant prp43p-TAP at NPT. The quantitation for fold change in U2, U5a, and U6 snRNAs associated with prp43p mutants is shown below the northern blot. Fold change was calculated as the percentage of each snRNA bound compared to wild-type. B). snRNAs associated with Prp43p-TAP and prp43L469Rp-TAP at 25° C or shifted to 37° C for 1 hour. U2, U6, and U5 snRNAs were visualized by northern blotting. Quantitation was preformed as described in (A). IN, 1% input; FT, 1% flow-through; B, 20% bound.

4.4 DISCUSSION

In this work, we characterized novel and previously identified alleles of *PRP43* with mutations in the DEAH-C. For each allele, we determine the effect of the mutant on ribosome biogenesis and pre-mRNA splicing. Using these alleles, we have related RNA binding and G-patch protein binding to specific regions Prp43p, which sheds light on its regulation. Of the mutants we analyzed, they fell into four different categories (Table A2). First, alleles *prp43G395R* and *prp43L469R* do not impair spliceosome recycling to the point that pre-mRNA splicing is stalled (Figs. 4.4 & 4.5), but each drastically affects ribosome biogenesis (Figs. 4.6 & 4.7). Though we did not observe a first-step splicing defect in *prp43G395R*, a substitution of glutamic acid at this position was previously reported to accumulate pre-mRNAs (Arenas and Abelson 1997). Since G395 lies near the base of the 5' HP, perhaps the difference in phenotype for these two mutants is related to a difference in the effect of the substitution on the movement of that structure. L469 lies at the interface between the WHD and the RecA1 domain. It is quite possible that the arginine substitution at this position interferes with the movement of the WHD domain that is predicted to occur during catalysis. Since L469 does not lie on the surface of the protein this change in the internal mechanism of the enzyme is most likely. However, interestingly, the conformational changes affected by G359R or L469R impact the ability of Prp43p to interact with two of its cofactors.

A second class of alleles with mutations in the CTD drastically affects both ribosome biogenesis and pre-mRNA splicing. Unlike the first class of mutants described, these CTD alleles have a strong defect in spliceosome recycling which results in a block

in the first step of splicing (Figs. 4.4 & 4.5). The triple alanine mutants ^{648}IRK^{650} and ^{717}WLI^{719} lie at the interface between two α helices distal to the OB-fold, and together line a small pocket (Fig. 2A right panel; Fig. 4.3B). The α helix containing ^{717}WLI^{719} also interfaces with the C-terminal α helix that is truncated in the *prp43Q732STOP* allele. This C-terminal α helix projects out into solution, but partially protects the surface it forms with L718 and I719. While mutations in L469 and G395 may directly impact conformational changes of Prp43p, mutations in the CTD are unlikely to directly affect catalysis. Though a portion of the CTD that encompasses part of the fourth β strand of the OB-fold and the loop between it and the fifth β sheet lies close to the 5' HP (amino acids 697-704), the other β strands of the OB-fold lie between this region and ^{648}IRK^{650} and ^{717}WLI^{719}. This loop is modeled to make contact with the RNA, and mutation of the loops between the fourth and fifth β sheets have been shown to reduce the RNA-stimulated ATPase activity of Prp43p (Walbott *et al*. 2010). It is possible that movement of the three α helices in which ^{648}IRK^{650} and ^{717}WLI^{719} lie and Q-STOP truncates indirectly affects the interaction between this loop and the RNA.

Our third class of mutants contains two alleles with two mutations each. While we cannot be certain at this time whether one or both of the mutations are required for the phenotype, it is curious that a mutation in each allele lies adjacent to one another, M492 and R493, implicating that region in the observed phenotype. Additionally, these two mutations lie along the interface between the WHD and the 5' HP. In particular, the mutation at R493 could impact 5' HP position; eleven of seventeen predicated rotamers

for this mutation clash with the tip of the 5' HP (R407). R493 lies one turn away on same α helix as E497, which is predicted to make a salt bridge with R407 (Walbott *et al*. 2010). These alleles are constitutively partially deficient in ribosome biogenesis with little impact on pre-mRNA splicing. Perhaps minor perturbations of the movement of the 5' HP are responsible for the phenotype observed. It is possible that ribosome biogenesis is more sensitive to the function of Prp43p. Association of none of the cofactors nor 35S rRNA were disrupted in these mutants, indicating these mutant proteins are being directed properly, but are partially incompetent mechanically.

A final mutant, *prp43A565E*, belongs in a separate class since growth is mildly affected, but no change in pre-mRNA splicing or ribosome biogenesis was observed. A565 lies in the ratchet domain, though in one of the α helices predicted to be part of the "motor" rather than the main α helix predicted to be the ratchet. We believe prp43A565Ep is being targeted properly, but is deficient in its ability to translocate. Further *in vitro* analysis will be needed to further understand how A565 affects Prp43p mechanistically.

Our data support a model in which more than one cofactor utilizes the same binding surface on Prp43p. It was previously reported that two mutations in the 5' HP of Prp43p cause temperature sensitivity, K398A and Y402A. The single mutation of Y402 results in loss of Ntr1p binding *in vitro*, and, moreover, lariat introns partially accumulate in post-spliceosomal complexes. Ntr1p forms one binding interface with this 5' HP region, which would be consistent with the model for activation of Prp43p by cofactors

(Tanaka *et al*. 2007; Walbott *et al*. 2010). Our data further support this model since the mutation of G395 also loses association with Ntr1p *in vivo*. As in the case of Y402, the mutation of G395 to arginine does not completely abolish the ability of Prp43p to function in pre-mRNA splicing. Our data along with data for Y402 indicate that cells can survive without Ntr1p/Prp43p interaction. Furthermore, mutation of G395 or L469 results in loss of Gno1p binding. We propose that both Ntr1p and Gno1p interact with the same regions of the 5' HP, and most likely are acting in similar manner to stimulate of the activity of Prp43p.

A single mutation in a conserved glycine of the G-patch domain of Ntr1p, L68A, results in its loss of interaction with Prp43p *in vitro*. While *ntr1L68A* is not a lethal allele, it is synthetically lethal with *prp43Y402A*. The *ntr1L68A* allele is also synthetically lethal in combination with a truncation of the CTD of Prp43p (aa 1-722) and triple alanine mutations of ^{717}WLI^{719}. However, the G-patch domain of Ntr1p makes contacts with Prp43p in the absence of the CTD (aa 1-712) *in vitro* (Tanaka *et al*.2007), proving that that region is not interacting with the CTD. These data were suggestive of a second function of the CTD of Prp43p, and more recent data have proven that the CTD is necessary for RNA binding *in vitro* (Walbott *et al*. 2010). Our data indicate that loss of RNA binding by the CTD is concomitant with loss of cofactor binding *in vivo*. Triple alanine mutants ^{648}IRK^{650} and ^{717}WLI^{719} lose interaction with Ntr1p and Gno1p, as well as snRNAs and the 35S pre-rRNA. While the Ntr1p/Prp43p interaction is dependent upon the G-patch domain of Ntr1p, it remains to be proven if Gno1p interacts in a similar manner. The mutation of a conserved tryptophan in the G-patch domain of Gno1p,

W38S, is sufficient to produce the rRNA processing defects observed in a *gno1Δ* (Guglielemi and Werner 2002). This single mutation does contribute to the interaction of Gno1p with Prp43p, but is not sufficient to completely disrupt their interaction *in vivo* (Fig. 3.6B). Perhaps other mutations in the G-patch domain will further elucidate the importance of the G-patch domain of Gno1p in relation to Prp43p. Nevertheless, we note that both Ntr1p and Gno1p require G-patch domains for their function *in vivo*, while the G-patch domain of Pfa1p is dispensable for growth and interaction with Prp43p (Tanaka *et al*. 2007; Guglielmi and Werner 2002; Pandit *et al*. 2009; Pertschy *et al*. 2009). Also, the G-patch domain of Pfa1p interacts with the C-terminus of Prp43p, not the β hairpin (Walbott *et al*. 2010). Because of the nonessential nature of Pfa1p, our data indicate that the effect on RNA processing observed in each mutant is unlikely to be linked to Pfa1p. It is possible that the dominant nature of Pfa1p in its association with Prp43p over the other two cofactors exacerbates weakened Prp43p/Ntr1p or Prp43p/Gno1p associations (Hennigan and Stevens, unpublished). However, there was not an observable increase in the amount of Prp43p/Pfa1p complexes in our mutants.

The dominant negative mutant prp43T123Ap retains the ability to bind the spliceosome, but is ATPase deficient and unable to unwind the intron lariat from post-splicing lariat RNP complexes (Martin *et al*. 2002; Tanaka and Schwer 2006). Interestingly, if the triple alanine mutations, IRK or WLI, are also introduced into *prp43T123A*, the dominant negative phenotype is relieved (Tanaka and Schwer *et al*. 2007). Therefore, these mutations in the CTD are affecting the targeting and binding of Prp43p to its substrates, which can relieve the dominance exhibited by the T123A

mutation. Since IRK and WLI both lose association with RNAs from both processes it is likely that this region in fact generally binds RNAs nonspecifically *in vivo*. Perhaps without the ability to bind RNA, the interaction between Prp43p and Ntr1p or Gno1p is not stabilized, which is why the loss of both RNA and cofactor binding is concomitant.

Chapter 5: Conclusions and Implications

5.1 Regulation of Prp43p via its interactions with G-patch proteins in yeast

In this work, we further define the previously suggested competition between the G-patch proteins (Pandit *et al*. 2009). We show that excess Pfa1p inhibits both the pre-mRNA splicing (Fig. 2.12) and ribosome biogenesis (Fig. 2.7) pathways directly by outcompeting both Ntr1p and Gno1p for Prp43p binding (Figs. 2.9 & 2.13). Under conditions in which Pfa1p levels are increased, Prp43p is inhibited from associating with either the 35S rRNA for ribosome biogenesis or the U6 snRNA in pre-mRNA splicing (Fig. 2.8). Additionally, we show that Ntr1p and Gno1p can compete with one another for Prp43p occupancy (Figs. 3.5 & 3.8). The overexpression of both Gno1p-HA/ZZ and Ntr1p-HA/ZZ results in mild effects on pre-mRNA splicing (Fig. 3.2) or pre-rRNA processing (Fig. 3.8A), respectively. The competition between Gno1p and Ntr1p can be visualized through the change in Prp43p localization upon their overexpression (Fig. 3.4). We note that Pfa1p is thought to be both nucleolar and nucleoplasmic, and its association with Prp43p does not change upon Ntr1p or Gno1p overexpression (Fig. 3.9). These data support a model in which Pfa1p has a dominant interaction with Prp43p over the other two cofactors. Since we used an *in vivo* system, Prp43p associations with the G-patch proteins were determined in the context of native complexes. *In vitro* competition assays and kinetic data are necessary to further define and clarify the direct relationships between the Prp43p and its cofactors. However, to accomplish these experiments the reported difficulties in properly expressing Gno1p must be overcome (Lebaron *et al*. 2009).

We believe that in response to stress, cells may use a regulatory mechanism that is dependent on Pfa1p (Fig. 2.14) to manipulate Prp43p distribution, thereby pausing both pre-rRNA and pre-mRNA processing pathways. This mechanism appears to be more complicated than simply increasing whole cell Pfa1p protein levels (Fig. 2.15), and will require further investigation. Our data also suggest that in response to heat shock, Pfa1p may be post-translationally modified (Fig. 2.15C). We are working to verify that such a modification exists and is functionally important. It will be interesting to determine if Pfa1p post-translational modification impacts Prp43p association.

Furthermore, we found that the G-patch domain of Gno1p is necessary for its interaction with Prp43p (Fig 3.6B), similar to that of Ntr1p (Tanaka *et al*. 2007), but in contrast to Pfa1p (Pandit *et al*. 2009; Lebaron *et al*. 2009). This interaction between the G-patch domain of Gno1p and Prp43p is also important for the association of Prp43p with the 35S pre-rRNA (Fig. 3.6C). These data suggest that the rRNA defects observed in the absence of Gno1p are most likely due to the loss of a Prp43p-mediated event. More biochemical data will be necessary to prove whether Gno1p binds Prp43p prior to pre-rRNA association or works to stabilize that association. Additionally, *in vitro* work with gno1W38Sp will be needed to demonstrate the loss of direct interaction and subsequent loss of Prp43p stimulation by Gno1p. Moreover, the exact function of Prp43p in pre-90S particles has not been determined. We note that there is an increase in the amount of 35S pre-rRNA in the *gno1Δ* strain, suggesting that 90S particles are stalled prior to Prp43p/Gno1p action. The characterization of the protein composition and rRNA structure of these stalled particles could help elucidate the role of Prp43p in the pre-90S.

The similarity of the interactions between Gno1p and Ntr1p with Prp43p suggests that they function with Prp43p in a similar manner. In contrast, Pfa1p appears to be unique in its interaction with Prp43p. Our analysis of *prp43* mutants further demonstrates this difference since multiple mutants simultaneously lose interaction with Ntr1p and Gno1p, but not Pfa1p (Figs. 4.10, 4.11 & 4.12). Previous work has implicated the C-terminus of Prp43p in pre-mRNA splicing (Tanaka *et al*. 2007), but not ribosome biogenesis. Our analysis of mutations in the DEAH-C of Prp43p has identified point mutants that strongly affect both pre-mRNA splicing and pre-rRNA processing. Both the *prp43IRK* and *prp43WLI* alleles lose snRNA and rRNA association (Figs. 4.13 & 4.14), which is consistent with the model that the OB-fold is needed for RNA binding (Walbott *et al*. 2010). Interestingly, these mutants lose interaction with Gno1p as well as with Ntr1p. (Figs. 4.11 & 4.12). This implies that the interaction surface of Prp43p is shared between Gno1p and Ntr1p. The loss of RNA interactions might contribute to the inability of Prp43p to incorporate into RNPs. Deletion of the OB-fold in the human ortholog of Prp43p (hPrp43p) resulted in a homogeneous localization of hPrp43p in the nucleus rather than in nuclear speckles or nucleoli (Fouraux *et al*. 2002). Our microscopy data demonstrating the delocalization of Prp43p when either Ntr1p/Prp43p or Gno1p/Prp43p interactions are disrupted (Fig. 3.4) supports the importance of the CTD in the subnuclear localization of Prp43p. Therefore, the ability of Prp43p to bind RNA, and therefore the basis for its specificity, is coupled to cofactor binding and stimulation. Once again, *in vitro* work with our *prp43* mutants will be instructive in illuminating their effects. In particular, we hypothesize that the growth phenotype observed in *prp43A565E* (Fig.

4.2B), which is mutated in the ratchet domain (Fig. 4.2A & 4.3A), is a result of a defect in translocation. *In vitro* duplex unwinding assays with the A565E mutant and other ratchet mutants will be necessary to determine their affect on the translocation and processivity of Prp43p. Mutants that lose interaction with the cofactors *in vivo* would be predicted to lose the stimulation of their ATPase activity. To fully rule out an effect of intrinsic ATPase activity in each *prp43* mutant, *in vitro* ATPase assays would be needed.

5.2 HUMAN PRP43P AND ITS INTERACTIONS WITH COFACTORS ARE CONSERVED

The human ortholog of Prp43p shares homology with the majority of the ScPrp43p, but contains a unique N-terminal region. Amino acids 25-60 consist of alternating positive and negative residues (Arg/Lys- Asp/Glu), and this region of hPrp43p is necessary for its proper localization. hPrp43p colocalizes to nuclear speckles with Sm proteins and U2B'' as well as in nucleoli (Fouraux *et al*. 2002), suggesting the dual function of Prp43p is conserved. However, it has yet to be proven that hPrp43p has a role in ribosome biogenesis. Establishing the role of hPrp43p in both ribosome biogenesis would be key to translating our data from yeast into humans.

In higher eukaryotes, nuclear speckles have been shown to be a storage site for splicing factors and the site of spliceosome assembly (Sleeman and Lamond 1999). In addition to the colocalization of hPrp43 with splicing factors, it associates with U1, U2, U4, U5 and U6 snRNAs. An N-terminal truncation of the first 62 amino acids of hPrp43p was shown to abrogate its localization to nucleoli, but not its localization to the nuclear speckles (Fouraux *et al*. 2002). Furthermore, an N-terminal truncation of the first 152 amino acids resulted in the cytoplasmic distribution of hPrp43p, indicating its NLS signal

lies within this region (Fouraux *et al*. 2002). Since this region is not conserved from the yeast homolog, the NLS signal appears to be divergent in yeast.

Additionally, hPrp43p does not share a twelve amino acid lysine rich C-terminal extension with ScPrp43p (Fouraux *et al*. 2002). A C-terminal truncation of 106 amino acids of hPrp43p also resulted in loss of its localization to nucleoli or nuclear speckles, but a homogeneous nuclear signal is observed. This C-terminal region is predicted to be necessary for targeting Prp43p to the subnuclear compartments accordingly (Fouraux *et al*. 2002). Interestingly, this C-terminal truncation of hPrp43p is upstream of the corresponding ^{648}IRK^{650} in ScPrp43p. This truncation would be missing part of the OB-fold and would be hypothesized to also lose interaction with RNA based on a similar truncation in ScPrp43p (Walbott *et al*. 2010). Our data demonstrates that prp43IRKp and prp43WLIp concomitantly lose interaction with RNAs (Fig. 4.13 & 4.14) and the pre-mRNA splicing or pre-ribosomal processing machineries. Our data along with the human data suggest that the contribution of RNA binding to the proper localization of Prp43p is conserved. Investigations of the unique domains of hPrp43p described above will be most interesting in determining any new roles of hPrp43p and its regulation in humans.

Interestingly, hPrp43p has been purified as a part of the 17S U2 snRNP, which is the species that joins the pre-mRNA (Hartmuth *et al*. 2002; Das *et al*. 2000). Our lab has been unable to demonstrate Prp43p is a component of the U2 snRNP in yeast (Hennigan and Stevens, unpublished). Two G-patch containing proteins are also found as part of purified human 17S U2 snRNPs, CHERP and SPF45 (Hartmuth *et al*. 2002). In *Drosophila* SPF45 (also termed RBM17) has been shown to play a role in the selection of

the 3' SS in alternative splicing, which has important implications in alternative splicing (Lallena *et al*. 2002). CHERP has not been shown to have a role in pre-mRNA splicing, but has a role in calcium homeostasis (Laplante *et al*. 2000). Neither is known to have a functional connection to Prp43p or any other DExH/D helicase. However, several other G-patch proteins have been shown to function with hPrp43p. Two of the yeast cofactors, Gno1p and Ntr1p, have obvious homologs in humans (Tannukit *et al*. 2009; Zhou and Lu 2001), and hPrp43p interacts with at least two other G-patch proteins, RBM5 and GPATCH2, not found in yeast (Lin *et al*. 2009; Niu *et al*. 2012).

The human homolog of the spliceosome related protein Ntr1p is TFIP11, which possesses 30% amino acid conservation (Tannukit *et al*. 2008). TFIP11 localizes to nuclear puncta, distinct from the nuclear speckles that contain splicing factors, and is excluded from nucleoli (Wen *et al*. 2005). The G-patch domain of TFIP11 is dispensable for its localization (Wen *et al*. 2008), but the N-terminal region of TFIP11 containing the G-patch domain is necessary its interaction with hPrp43p (Yoshimoto *et al*. 2009). As in yeast, TFIP11/hPrp43p association is necessary to disassemble the post-splicing lariat RNP, indicating that they function together for the same role in pre-mRNA splicing in humans (Yoshimoto *et al*. 2009). Further investigations will be necessary to determine if TFIP11 acts differently in alternative splicing scenarios.

hPINX1 is the human homolog of yeast protein Gno1p with approximately 50% amino acid similarity. However, though it localizes to the nucleolus, hPINX1 has yet to be proven to function in ribosome biogenesis (Zhou and Lu 2001). Despite this lack of data, hPINX1 rescues growth and ribosome biogenesis defects when expressed in a

gno1Δ strain, indicating that hPINX1 can support the role Gno1p plays in rRNA processing, at least in yeast (Guglielmi and Werner 2002). These data suggest that hPINX1 likely plays a role in human rRNA processing, but further work will be necessary to demonstrate a functional relationship between hPrp43p and hPINX1.

hPINX1 has an additional function in maintaining telomeres by acting as an inhibitor of telomerase (Zhou and Lu 2001). hPinX1 binds directly to the telomerase catalytic component TERT (Zhou and Lu 2001), and recruits and mediates the localization of human telomerase to the telomere in a cell cycle-dependent manner (Cheung *et al*. 2012). In yeast, Gno1p competitively binds the same interface of TERT (Est2p) as the RNA component of telomerase (TLC1). While bound to Gno1p, TERT is sequestered from TLC1 and remains in an inactive state. While the G-patch domain of Gno1p is necessary for its role in ribosome biogenesis, it is dispensable for its function in telomere maintenance (Lin and Blackburn 2004; Guglielmi and Werner 2002). *In vitro* work will be needed to demonstrate the relationship between the G-patch domain of hPINX1 in rRNA processing and hPrp43p.

The G-patch protein RBM5 has recently been shown to interact with hPrp43 *in vivo* in HeLa cells and directly *in vitro*. This interaction is mediated by the C-terminus of RBM5, which contains an essential G-patch domain (aa 742-815). It has already been shown that two or more point mutations of the glycines in this domain strongly reduce the binding of RBM5 to hPrp43, and, subsequently, the ATPase and helicase activity of hPrp43p is no longer stimulated by RBM5 (Niu *et al*. 2012). RBM5 was not detected in post-splicing lariat RNPs, so it is unlikely to be working in disassembly (Yoshimoto *et al*.

2009). Therefore, it is still unknown in which process RBM5 and hPrp43p are functioning together.

RBM5 has also been shown to be a tumor suppressor that induces cell cycle arrest and apoptosis through cyclin A or Bax pathways, respectively, by regulating the splicing of pre-mRNA of target genes in those pathways (Oh *et al*. 2006). An Octamer Repeat (OCRE) domain in RBM5 contacts U5 and is necessary to modulate alternative splicing after assembly and exon definition (Bonnal *et al*. 2008). RBM5 is one of 35 genes related to tumor suppression in a 370 kB region on chromosome 3p21.3. The loss of this region is the most frequent and earliest event in the development of non-small cell lung cancer (Wei *et al*. 1996). In non-small cell lung cancers, RBM5 mRNA and protein levels were shown to be significantly reduced in patient tissue samples. Additionally, the decrease in RBM5 was positively correlated to lymph node metastasis and also seen in IIIA and IIIB stage cancers (Liang *et al*. 2012). Similarly, the mRNA and protein levels of RBM5 are downregulated in pancreatic cancer tissue compared to non-tumor tissue. This reduced level in RBM5 also correlates with in an increase in metastasis (Peng *et al*. 2013). These links between RBM5 and disease highlight the importance in understanding its relationship with hPrp43p.

Finally, very little is known about the G-patch containing protein GPATCH2, which has been shown to interact with hPrp43p in beast carcinomas in a single report (Lin *et al*. 2009). This novel protein was identified as overexpressed by 3 fold in about 40% of breast cancers tested. While, GPATCH2 has obvious murine and Xenopus homologues (Lin *et al*. 2009), there is no identified yeast homolog. hPrp43p

coimmunoprecipitates with GPATCH2 and colocalizes in both nuclear speckles and nucleoli. When protein levels of hPrp43p are decreased via siRNA treatment, breast cancer cells have reduced viability. Interestingly, the ATPase activity of recombinant hPrp43p was increased by 45-fold in the presence of GPATCH2, implicating it as a cofactor in an unknown pathway. The stimulation of hPrp43p by GPATCH2 could contribute to the transformation of normal beast cells, and further investigation into the nature and implications of this interaction is warranted for basis of further therapeutic study (Lin *et al*. 2009).

In conclusion, potentially the regulation of hPrp43p by G-patch proteins is conserved in humans. Much work must be done with each of the four cofactors to understand their relationships with hPrp43p and roles in humans. Additionally, in humans, at least 22 G-patch domain-containing proteins are predicted, more than 4 times the number in yeast (The UniProt Consortium 2011). Likewise, the number of DEAH-box helicases predicted in humans is 14, double the number identified in yeast (Umate *et al*. 2011). If the function of all of the human G-patch proteins is in regulating DEAH-box proteins, then the complexity of the regulation is even greater than that in yeast. It will be interesting to determine how other G-patch-containing proteins in metazoa are involved in regulating the activity of additional helicases by stimulation and/or recruitment.

Appendix

Table A1. All *prp43* alleles isolated from gap-repair mutagenesis in Chapter 4.

Random *prp43* alleles	Phenotype
L469R D567V	ts
E451D M492T	ts
F437L A502P	ts
G473D R493K	ts
L718P F656S	ts
Y440C	ts
R370G I619T	ts
A565E	ts
E10K, E16Q, L698P	ts
P127S, G434A, E497G, S702F	ts
E67K, E117K, P268S	ts
S59T, P73S, E11K, E134K, P454S, T559A, K745R	ts
H298Y, D386G, S621F, M658T, S665L	ts
H298Y, E443K, D592N, W717R	ts
N618I, H682R, Q732STOP	ts
D638N	cs

Table A2. Summary of characteristics for *prp43* alleles.

Mutant	Amino Acid Change	snRNP levels	pre-mRNA accumulation	Ribosome Biogenesis defects	35S association	snRNA association	Cofactor associations lost
prp43IRK	IRK 648-650 AAA	↑U4/U6 ↓tri-snRNP	↑	↓ 20S ↓ 27S ↑ 35S	Reduced	Reduced	Gno1p, Ntr1p
prp43WLI	WLI 717-719 AAA	↑U4/U6 ↓tri-snRNP	slightly↑	↓ 20S ↓ 27S ↑ 35S	Reduced	Reduced	Gno1p, Ntr1p
prp43Q732STOP	CΔ35	↑U4/U6 ↓tri-snRNP	↑	↓ 20S ↓ 27S ↑ 35S	N/A	N/A	N/A
prp43G359R	G395R	↑U4/U6 slightly ↓tri-snRNP intermediate to WT	None	↓ 27S ↑ 35S	Reduced	Reduced	Gno1p, Ntr1p
prp43L469R	L469R	↑U4/U6 slightly ↓tri-snRNP intermediate to WT	None	↓ 20S ↓ 27S ↓ 35S	Reduced	Reduced	Gno1p, Ntr1p
prp43-1504	G473D R493K	WT	None	↑ 35S overall rRNAs decreased	WT	WT	None
prp43-1248	E451D M492T	WT	None	↑ 35S overall rRNAs decreased	WT	WT	None
prp43A565E	A565E	WT	None	WT			

References

Abdelhaleem M. 2004. Do human RNA helicases have a role in cancer? *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1704**: 37–46.

Achsel T, Brahms H, Kastner B, Bachi A, Wilm M, Lührmann R. 1999. A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J* **18**: 5789–5802.

Anantharaman V, Koonin EV, Aravind L. 2002. Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucleic Acids Res* **30**: 1427–1464.

Anthony JG, Weidenhammer EM, Woolford JL. 1997. The yeast Prp3 protein is a U4/U6 snRNP protein necessary for integrity of the U4/U6 snRNP and the U4/U6.U5 tri-snRNP. *RNA* **3**: 1143–1152.

Aravind L, Koonin EV. 1999. G-patch: a new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins. *Trends in Biochemical Sciences* **24**: 342–344.

Arenas JE, Abelson JN. 1997. Prp43: An RNA helicase-like factor involved in spliceosome disassembly. *Proc Natl Acad Sci U S A* **94**: 11798–11802.

Ares M, Grate L, Pauling MH. 1999. A handful of intron-containing genes produces the lion's share of yeast mRNA. *RNA* **5**: 1138–1139.

Armakola M, Higgins MJ, Figley MD, Barmada SJ, Scarborough EA, Diaz Z, Fang X, Shorter J, Krogan NJ, Finkbeiner S, et al. 2012. Inhibition of RNA lariat debranching enzyme suppresses TDP-43 toxicity in ALS disease models. *Nature Genetics* **44**: 1302–1309.

Ben-Aroya S, Coombes C, Kwok T, O'Donnell KA, Boeke JD, Hieter P. 2008. Toward a comprehensive temperature-sensitive mutant repository of the essential genes of Saccharomyces cerevisiae. *Mol Cell* **30**: 248–258.

Azubel M, Habib N, Sperling R, Sperling J. 2006. Native Spliceosomes Assemble with Pre-mRNA to Form Supraspliceosomes. *Journal of Molecular Biology* **356**: 955–966.

Banroques J, Abelson JN. 1989. PRP4: a protein of the yeast U4/U6 small nuclear ribonucleoprotein particle. *Mol Cell Biol* **9**: 3710–3719.

Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Goffeau A, Hoheisel JD, Jacq C, Johnston M, et al. 1996. Life with 6000 genes. *Science* **274**: 546+.

Bergkessel M, Whitworth GB, Guthrie C. 2011. Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast. *RNA* **17**: 1461–1478.

Bernstein KA, Granneman S, Lee AV, Manickam S, Baserga SJ. 2006. Comprehensive mutational analysis of yeast DEXD/H box RNA helicases involved in large ribosomal subunit biogenesis. *Mol Cell Biol* **26**: 1195–1208.

Berry DB, Gasch AP. 2008. Stress-activated genomic expression changes serve a preparative role for impending stress in yeast. *Mol Biol Cell* **19**: 4580–4587.

Bohnsack MT, Kos M, Tollervey D. 2008. Quantitative analysis of snoRNA association with pre-ribosomes and release of snR30 by Rok1 helicase. *EMBO Rep* **9**: 1230–1236.

Bohnsack MT, Martin R, Granneman S, Ruprecht M, Schleiff E, Tollervey D. 2009. Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. *Mol Cell* **36**: 583–592.

Bolger TA, Wente SR. 2011. Gle1 Is a Multifunctional DEAD-box Protein Regulator That Modulates Ded1 in Translation Initiation. *Journal of Biological Chemistry* **286**: 39750–39759.

Bonnal S, Martínez C, Förch P, Bachi A, Wilm M, Valcárcel J. 2008. RBM5/Luca-15/H37 Regulates Fas Alternative Splice Site Pairing after Exon Definition. *Molecular Cell* **32**: 81–95.

Boon K-L, Auchynnikava T, Edwalds-Gilbert G, Barrass JD, Droop AP, Dez C, Beggs JD. 2006. Yeast ntr1/spp382 mediates prp43 function in postspliceosomes. *Mol Cell Biol* **26**: 6016–6023.

Boon K-L, Grainger RJ, Ehsani P, Barrass JD, Auchynnikava T, Inglehearn CF, Beggs JD. 2007. prp8 mutations that cause human retinitis pigmentosa lead to a U5 snRNP maturation defect in yeast. *Nat Struct Mol Biol* **14**: 1077–1083.

Brand RC, Klootwijk J, Van Steenbergen TJ, De Kok AJ, Planta RJ. 1977. Secondary methylation of yeast ribosomal precursor RNA. *Eur J Biochem* **75**: 311–318.

Brauer MJ, Huttenhower C, Airoldi EM, Rosenstein R, Matese JC, Gresham D, Boer VM, Troyanskaya OG, Botstein D. 2008. Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. *Mol Biol Cell* **19**: 352–367.

Breathnach R, Benoist C, O'Hare K, Gannon F, Chambon P. 1978. Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *PNAS* **75**: 4853–4857.

Brody E, Abelson J. 1985. The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science* **228**: 963–967.

Bromley S, Hereford L, Rosbash M. 1982. Further evidence that the rna2 mutation of Saccharomyces cerevisiae affects mRNA processing. *Mol Cell Biol* **2**: 1205–1211.

Brow DA, Guthrie C. 1988. Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *Nature* **334**: 213–218.

Brown JWS, Marshall DF, Echeverria M. 2008. Intronic noncoding RNAs and splicing. *Trends in Plant Science* **13**: 335–342.

Burgess S, Couto JR, Guthrie C. 1990. A putative ATP binding protein influences the fidelity of branchpoint recognition in yeast splicing. *Cell* **60**: 705–717.

Burgess SM, Guthrie C. 1993. Beat the clock: paradigms for NTPases in the maintenance of biological fidelity. *Trends Biochem Sci* **18**: 381–384.

Bussiere C, Hashem Y, Arora S, Frank J, Johnson AW. 2012. Integrity of the P-site is probed during maturation of the 60S ribosomal subunit. *J Cell Biol* **197**: 747–759.

Büttner K, Nehring S, Hopfner K-P. 2007a. Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nature Structural & Molecular Biology* **14**: 647–652.

Cannone JJ, Subramanian S, Schnare MN, Collett JR, D'Souza LM, Du Y, Feng B, Lin N, Madabusi LV, Müller KM, et al. 2002. The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* **3**: 2.

Chakarova CF, Hims MM, Bolz H, Abu-Safieh L, Patel RJ, Papaioannou MG, Inglehearn CF, Keen TJ, Willis C, Moore AT, et al. 2002. Mutations in HPRP3, a third member of pre-mRNA splicing factor genes, implicated in autosomal dominant retinitis pigmentosa. *Hum Mol Genet* **11**: 87–92.

Chamberlain JR, Lee Y, Lane WS, Engelke DR. 1998. Purification and characterization of the nuclear RNase P holoenzyme complex reveals extensive subunit overlap with RNase MRP. *Genes Dev* **12**: 1678–1690.

Chan S-P, Cheng S-C. 2005. The Prp19-associated Complex Is Required for Specifying Interactions of U5 and U6 with Pre-mRNA during Spliceosome Activation. *J Biol Chem* **280**: 31190–31199.

Chapman KB, Boeke JD. 1991. Isolation and characterization of the gene encoding yeast debranching enzyme. *Cell* **65**: 483–492.

Chen H-C, Tseng C-K, Tsai R-T, Chung C-S, Cheng S-C. 2013. Link of NTR-Mediated Spliceosome Disassembly with DEAH-Box ATPases Prp2, Prp16, and Prp22. *Mol Cell Biol* **33**: 514–525.

Chen JH, Lin RJ. 1990. The yeast PRP2 protein, a putative RNA-dependent ATPase, shares extensive sequence homology with two other pre-mRNA splicing factors. *Nucleic Acids Res* **18**: 6447.

Chen JY-F, Stands L, Staley JP, Jackups Jr. RR, Latus LJ, Chang T-H. 2001. Specific Alterations of U1-C Protein or U1 Small Nuclear RNA Can Eliminate the Requirement of Prp28p, an Essential DEAD Box Splicing Factor. *Molecular Cell* **7**: 227–232.

Chen Y-IG, Moore RE, Ge HY, Young MK, Lee TD, Stevens SW. 2007. Proteomic analysis of in vivo-assembled pre-mRNA splicing complexes expands the catalog of participating factors. *Nucleic Acids Res* **35**: 3928–3944.

Cheng Z, Menees TM. 2004. RNA branching and debranching in the yeast retrovirus-like element Ty1. *Science* **303**: 240–243.

Cheung DH-C, Kung H-F, Huang J-J, Shaw P-C. 2012. PinX1 is involved in telomerase recruitment and regulates telomerase function by mediating its localization. *FEBS Lett* **586**: 3166–3171.

Choesmel V, Fribourg S, Aguissa-Touré A-H, Pinaud N, Legrand P, Gazda HT, Gleizes P-E. 2008. Mutation of ribosomal protein RPS24 in Diamond-Blackfan anemia results in a ribosome biogenesis disorder. *Hum Mol Genet* **17**: 1253–1263.

Colau G, Thiry M, Leduc V, Bordonné R, Lafontaine DLJ. 2004. The small nucle(ol)ar RNA cap trimethyltransferase is required for ribosome synthesis and intact nucleolar morphology. *Mol Cell Biol* **24**: 7976–7986.

Colley A, Beggs JD, Tollervey D, Lafontaine DL. 2000. Dhr1p, a putative DEAH-box RNA helicase, is associated with the box C+D snoRNP U3. *Mol Cell Biol* **20**: 7238–7246.

Combs DJ, Nagel RJ, Ares M Jr, Stevens SW. 2006. Prp43p is a DEAH-box spliceosome disassembly factor essential for ribosome biogenesis. *Mol Cell Biol* **26**: 523–534.

Company M, Arenas J, Abelson J. 1991. Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. *Nature* **349**: 487–493.

Cooper TA, Wan L, Dreyfuss G. 2009. RNA and Disease. *Cell* **136**: 777–793.

Cordin O, Hahn D, Beggs JD. 2012. Structure, function and regulation of spliceosomal RNA helicases. *Current Opinion in Cell Biology* **24**: 431–438.

Das R, Zhou Z, Reed R. 2000. Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. *Mol Cell* **5**: 779–787.

Dhote V, Sweeney TR, Kim N, Hellen CUT, Pestova TV. 2012. Roles of individual domains in the function of DHX29, an essential factor required for translation of structured mammalian mRNAs. *Proc Natl Acad Sci USA* **109**: E3150–3159.

Domdey H, Apostol B, Lin RJ, Newman A, Brody E, Abelson J. 1984. Lariat structures are in vivo intermediates in yeast pre-mRNA splicing. *Cell* **39**: 611–621.

Dragon F, Gallagher JEG, Compagnone-Post PA, Mitchell BM, Porwancher KA, Wehner KA, Wormsley S, Settlage RE, Shabanowitz J, Osheim Y, et al. 2002. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* **417**: 967–970.

Draptchinskaia N, Gustavsson P, Andersson B, Pettersson M, Willig TN, Dianzani I, Ball S, Tchernia G, Klar J, Matsson H, et al. 1999. The gene encoding ribosomal protein S19 is mutated in Diamond-Blackfan anaemia. *Nat Genet* **21**: 169–175.

Dunbar DA, Baserga SJ. 1998. The U14 snoRNA is required for 2'-O-methylation of the pre-18S rRNA in Xenopus oocytes. *RNA* **4**: 195–204.

Dziembowski A, Ventura A-P, Rutz B, Caspary F, Faux C, Halgand F, Laprévote O, Séraphin B. 2004. Proteomic analysis identifies a new complex required for nuclear pre-mRNA retention and splicing. *EMBO J* **23**: 4847–4856.

Edery P, Marcaillou C, Sahbatou M, Labalme A, Chastang J, Touraine R, Tubacher E, Senni F, Bober MB, Nampoothiri S, et al. 2011. Association of TALS Developmental Disorder with Defect in Minor Splicing Component U4atac snRNA. *Science* **332**: 240–243.

Emery B, De La Cruz J, Rocak S, Deloche O, Linder P. 2004. Has1p, a member of the DEAD-box family, is required for 40S ribosomal subunit biogenesis in Saccharomyces cerevisiae†. *Molecular Microbiology* **52**: 141–158.

Fabrizio P, Dannenberg J, Dube P, Kastner B, Stark H, Urlaub H, Lührmann R. 2009. The Evolutionarily Conserved Core Design of the Catalytic Activation Step of the Yeast Spliceosome. *Molecular Cell* **36**: 593–608.

Fairman-Williams ME, Guenther U-P, Jankowsky E. 2010. SF1 and SF2 helicases: family matters. *Curr Opin Struct Biol* **20**: 313–324.

Fayet-Lebaron E, Atzorn V, Henry Y, Kiss T. 2009. 18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences. *EMBO J* **28**: 1260–1270.

Fortner DM, Troy RG, Brow DA. 1994. A stem/loop in U6 RNA defines a conformational switch required for pre-mRNA splicing. *Genes Dev* **8**: 221–233.

Fouraux MA, Kolkman MJM, Van der Heijden A, De Jong AS, Van Venrooij WJ, Pruijn GJM. 2002. The human La (SS-B) autoantigen interacts with DDX15/hPrp43, a putative DEAH-box RNA helicase. *RNA* **8**: 1428–1443.

Fourmann J-B, Schmitzová J, Christian H, Urlaub H, Ficner R, Boon K-L, Fabrizio P, Lührmann R. 2013. Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. *Genes Dev* **27**: 413–428.

Freed EF, Bleichert F, Dutca LM, Baserga SJ. 2010. When ribosomes go bad: diseases of ribosome biogenesis. *Mol Biosyst* **6**: 481–493.

Galisson F, Legrain P. 1993. The biochemical defects of prp4-1 and prp6-1 yeast splicing mutants reveal that the PRP6 protein is required for the accumulation of the [U4/U6.U5] tri-snRNP. *Nucleic Acids Res* **21**: 1555–1562.

Gallagher JEG, Dunbar DA, Granneman S, Mitchell BM, Osheim Y, Beyer AL, Baserga SJ. 2004. RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components. *Genes Dev* **18**: 2506–2517.

Ganot P, Jády BE, Bortolin ML, Darzacq X, Kiss T. 1999. Nucleolar factors direct the 2'-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. *Mol Cell Biol* **19**: 6906–6917.

Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* **11**: 4241–4257.

Gee S, Krauss SW, Miller E, Aoyagi K, Arenas J, Conboy JG. 1997. Cloning of mDEAH9, a putative RNA helicase and mammalian homologue of Saccharomyces cerevisiae splicing factor Prp43. *PNAS* **94**: 11803–11807.

Gietz RD, Woods RA. 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Meth Enzymol* **350**: 87–96.

Gottschalk A, Kastner B, Lührmann R, Fabrizio P. 2001. The yeast U5 snRNP coisolated with the U1 snRNP has an unexpected protein composition and includes the splicing factor Aar2p. *RNA* **7**: 1554–1565.

Gozani O, Feld R, Reed R. 1996. Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. *Genes Dev* **10**: 233–243.

Grabowski PJ, Padgett RA, Sharp PA. 1984. Messenger RNA splicing in vitro: an excised intervening sequence and a potential intermediate. *Cell* **37**: 415–427.

Granneman S, Bernstein KA, Bleichert F, Baserga SJ. 2006. Comprehensive mutational analysis of yeast DEXD/H box RNA helicases required for small ribosomal subunit synthesis. *Mol Cell Biol* **26**: 1183–1194.

Guglielmi B, Werner M. 2002. The yeast homolog of human PinX1 is involved in rRNA and small nucleolar RNA maturation, not in telomere elongation inhibition. *J Biol Chem* **277**: 35712–35719.

Häcker I, Sander B, Golas MM, Wolf E, Karagöz E, Kastner B, Stark H, Fabrizio P, Lührmann R. 2008. Localization of Prp8, Brr2, Snu114 and U4/U6 proteins in the yeast tri-snRNP by electron microscopy. *Nature Structural & Molecular Biology* **15**: 1206–1212.

Hartmuth K, Urlaub H, Vornlocher H-P, Will CL, Gentzel M, Wilm M, Lührmann R. 2002. Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. *Proc Natl Acad Sci USA* **99**: 16719–16724.

Hasegawa M, Arai T, Akiyama H, Nonaka T, Mori H, Hashimoto T, Yamazaki M, Oyanagi K. 2007. TDP-43 is deposited in the Guam parkinsonism-dementia complex brains. *Brain* **130**: 1386–1394.

He H, Liyanarachchi S, Akagi K, Nagy R, Li J, Dietrich RC, Li W, Sebastian N, Wen B, Xin B, et al. 2011. Mutations in U4atac snRNA, a Component of the Minor Spliceosome, in the Developmental Disorder MOPD I. *Science* **332**: 238–240.

He Y, Andersen GR, Nielsen KH. 2010. Structural basis for the function of DEAH helicases. *EMBO Rep* **11**: 180–186.

Hegele A, Kamburov A, Grossmann A, Sourlis C, Wowro S, Weimann M, Will CL, Pena V, Lührmann R, Stelzl U. 2012. Dynamic Protein-Protein Interaction Wiring of the Human Spliceosome. *Molecular Cell* **45**: 567–580.

Henras AK, Soudet J, Gérus M, Lebaron S, Caizergues-Ferrer M, Mougin A, Henry Y. 2008. The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci* **65**: 2334–2359.

Henry Y, Wood H, Morrissey JP, Petfalski E, Kearsey S, Tollervey D. 1994. The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. *EMBO J* **13**: 2452–2463.

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**: 51–59.

Horn DM, Mason SL, Karbstein K. 2011. Rcl1 Protein, a Novel Nuclease for 18 S Ribosomal RNA Production. *J Biol Chem* **286**: 34082–34087.

Hoskins AA, Friedman LJ, Gallagher SS, Crawford DJ, Anderson EG, Wombacher R, Ramirez N, Cornish VW, Gelles J, Moore MJ. 2011. Ordered and dynamic assembly of single spliceosomes. *Science* **331**: 1289–1295.

Hoskins AA, Moore MJ. 2012. The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem Sci* **37**: 179–188.

Howson R, Huh W-K, Ghaemmaghami S, Falvo JV, Bower K, Belle A, Dephoure N, Wykoff DD, Weissman JS, O'Shea EK. 2005. Construction, verification and experimental use of two epitope-tagged collections of budding yeast strains. *Comp Funct Genomics* **6**: 2–16.

Huber J, Cronshagen U, Kadokura M, Marshallsay C, Wada T, Sekine M, Luhrmann R. 1998. Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *EMBO J* **17**: 4114–4126.

Hughes JM, Ares M Jr. 1991. Depletion of U3 small nucleolar RNA inhibits cleavage in the 5' external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA. *EMBO J* **10**: 4231–4239.

Huh W-K, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK. 2003. Global analysis of protein localization in budding yeast. *Nature* **425**: 686–691.

Idol RA, Robledo S, Du H-Y, Crimmins DL, Wilson DB, Ladenson JH, Bessler M, Mason PJ. 2007. Cells depleted for RPS19, a protein associated with Diamond Blackfan Anemia, show defects in 18S ribosomal RNA synthesis and small ribosomal subunit production. *Blood Cells, Molecules, and Diseases* **39**: 35–43.

Jády BE, Darzacq X, Tucker KE, Matera AG, Bertrand E, Kiss T. 2003. Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. *EMBO J* **22**: 1878–1888.

James S-A, Turner W, Schwer B. 2002. How Slu7 and Prp18 cooperate in the second step of yeast pre-mRNA splicing. *RNA* **8**: 1068–1077.

Jankowsky E. 2011. RNA helicases at work: binding and rearranging. *Trends Biochem Sci* **36**: 19–29.

Juneau K, Palm C, Miranda M, Davis RW. 2007. High-density yeast-tiling array reveals previously undiscovered introns and extensive regulation of meiotic splicing. *PNAS* **104**: 1522–1527.

Karni R, de Stanchina E, Lowe SW, Sinha R, Mu D, Krainer AR. 2007. The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nat Struct Mol Biol* **14**: 185–193.

Kastner B, Bach M, Lührmann R. 1990. Electron microscopy of small nuclear ribonucleoprotein (snRNP) particles U2 and U5: evidence for a common structure-determining principle in the major U snRNP family. *Proc Natl Acad Sci USA* **87**: 1710–1714.

Kim J-W, Kim H-C, Kim G-M, Yang J-M, Boeke JD, Nam K. 2000. Human RNA lariat debranching enzyme cDNA complements the phenotypes of Saccharomyces cerevisiae dbr1 and Schizosaccharomyces pombe dbr1 mutants. *Nucleic Acids Res* **28**: 3666–3673.

Kim SH, Smith J, Claude A, Lin RJ. 1992. The purified yeast pre-mRNA splicing factor PRP2 is an RNA-dependent NTPase. *EMBO J* **11**: 2319–2326.

King DS, Beggs JD. 1990. Interactions of PRP2 protein with pre-mRNA splicing complexes in Saccharomyces cerevisiae. *Nucleic Acids Res* **18**: 6559–6564.

King TH, Liu B, McCully RR, Fournier MJ. 2003. Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Mol Cell* **11**: 425–435.

Kiss T, Fayet E, Jády BE, Richard P, Weber M. 2006. Biogenesis and intranuclear trafficking of human box C/D and H/ACA RNPs. *Cold Spring Harb Symp Quant Biol* **71**: 407–417.

Kistler AL, Guthrie C. 2001. Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for Sub2, an essential spliceosomal ATPase. *Genes Dev* **15**: 42–49.

Konarska MM, Vilardell J, Query CC. 2006. Repositioning of the Reaction Intermediate within the Catalytic Center of the Spliceosome. *Molecular Cell* **21**: 543–553.

Koodathingal P, Novak T, Piccirilli JA, Staley JP. 2010. The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing. *Mol Cell* **39**: 385–395.

Kos M, Tollervey D. 2005. The Putative RNA Helicase Dbp4p Is Required for Release of the U14 snoRNA from Preribosomes in Saccharomyces cerevisiae. *Molecular Cell* **20**: 53–64.

Krawczak M, Reiss J, Cooper DN. 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum Genet* **90**: 41–54.

Kressler D, Hurt E, Baβler J. 2010. Driving ribosome assembly. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1803**: 673–683.

Kudlinzki D, Schmitt A, Christian H, Ficner R. 2012. Structural analysis of the C-terminal domain of the spliceosomal helicase Prp22. *Biol Chem* **393**: 1131–1140.

Kufel J, Dichtl B, Tollervey D. 1999. Yeast Rnt1p is required for cleavage of the pre-ribosomal RNA in the 3' ETS but not the 5' ETS. *RNA* **5**: 909–917.

Lafontaine DL, Preiss T, Tollervey D. 1998. Yeast 18S rRNA dimethylase Dim1p: a quality control mechanism in ribosome synthesis? *Mol Cell Biol* **18**: 2360–2370.

Lagier-Tourenne C, Polymenidou M, Cleveland DW. 2010. TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. *Hum Mol Genet* **19**: R46–R64.

Lallena MJ, Chalmers KJ, Llamazares S, Lamond AI, Valcárcel J. 2002. Splicing Regulation at the Second Catalytic Step by Sex-lethal Involves 3′ Splice Site Recognition by SPF45. *Cell* **109**: 285–296.

Lange TS, Gerbi SA. 2000. Transient nucleolar localization Of U6 small nuclear RNA in Xenopus Laevis oocytes. *Mol Biol Cell* **11**: 2419–2428.

Laplante JM, O'Rourke F, Lu X, Fein A, Olsen A, Feinstein MB. 2000. Cloning of human Ca2+ homoeostasis endoplasmic reticulum protein (CHERP): regulated expression of antisense cDNA depletes CHERP, inhibits intracellular Ca2+ mobilization and decreases cell proliferation. *Biochem J* **348 Pt 1**: 189–199.

Lardelli RM, Thompson JX, Yates JR, Stevens SW. 2010. Release of SF3 from the intron branchpoint activates the first step of pre-mRNA splicing. *RNA* **16**: 516–528.

Lebaron S, Froment C, Fromont-Racine M, Rain J-C, Monsarrat B, Caizergues-Ferrer M, Henry Y. 2005. The Splicing ATPase Prp43p Is a Component of Multiple Preribosomal Particles. *Mol Cell Biol* **25**: 9269–9282.

Lebaron S, Papin C, Capeyrou R, Chen Y-L, Froment C, Monsarrat B, Caizergues-Ferrer M, Grigoriev M, Henry Y. 2009. The ATPase and helicase activities of Prp43p are stimulated by the G-patch protein Pfa1p during yeast ribosome biogenesis. *EMBO J* **28**: 3808–3819.

Lebaron S, Schneider C, Nues RW van, Swiatkowska A, Walsh D, Böttcher B, Granneman S, Watkins NJ, Tollervey D. 2012. Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits. *Nature Structural & Molecular Biology* **19**: 744–753.

Lee I, Li Z, Marcotte EM. 2007. An improved, bias-reduced probabilistic functional gene network of baker's yeast, Saccharomyces cerevisiae. *PLoS ONE* **2**: e988.

Leeds NB, Small EC, Hiley SL, Hughes TR, Staley JP. 2006. The Splicing Factor Prp43p, a DEAH Box ATPase, Functions in Ribosome Biogenesis. *Mol Cell Biol* **26**: 513–522.

Legrain P, Seraphin B, Rosbash M. 1988. Early commitment of yeast pre-mRNA to the spliceosome pathway. *Mol Cell Biol* **8**: 3755–3760.

Lesser CF, Guthrie C. 1993. Mutations in U6 snRNA that alter splice site specificity: implications for the active site. *Science* **262**: 1982–1988.

Li HD, Zagorski J, Fournier MJ. 1990. Depletion of U14 small nuclear RNA (snR128) disrupts production of 18S rRNA in Saccharomyces cerevisiae. *Mol Cell Biol* **10**: 1145–1152.

Li Z, Brow DA. 1993. A rapid assay for quantitative detection of specific RNAs. *Nucleic Acids Res* **21**: 4645–4646.

Liang H, Zhang J, Shao C, Zhao L, Xu W, Sutherland LC, Wang K. 2012. Differential expression of RBM5, EGFR and KRAS mRNA and protein in non-small cell lung cancer tissues. *Journal of Experimental & Clinical Cancer Research* **31**: 36.

Liang WQ, Fournier MJ. 1995. U14 base-pairs with 18S rRNA: a novel snoRNA interaction required for rRNA processing. *Genes Dev* **9**: 2433–2443.

Liang X, Fournier MJ. 2006. The Helicase Has1p Is Required for snoRNA Release from Pre-rRNA. *Mol Cell Biol* **26**: 7437–7450.

Liang X, Liu Q, Fournier MJ. 2007. rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. *Mol Cell* **28**: 965–977.

Lin J, Blackburn EH. 2004. Nucleolar protein PinX1p regulates telomerase by sequestering its protein catalytic subunit in an inactive complex lacking telomerase RNA. *Genes Dev* **18**: 387–396.

Lin M-L, Fukukawa C, Park J-H, Naito K, Kijima K, Shimo A, Ajiro M, Nishidate T, Nakamura Y, Katagiri T. 2009. Involvement of G-patch domain containing 2 overexpression in breast carcinogenesis. *Cancer Sci* **100**: 1443–1450.

Lin RJ, Lustig AJ, Abelson J. 1987. Splicing of yeast nuclear pre-mRNA in vitro requires a functional 40S spliceosome and several extrinsic factors. *Genes Dev* **1**: 7–18.

Lo K-Y, Li Z, Bussiere C, Bresson S, Marcotte EM, Johnson AW. 2010. Defining the pathway of cytoplasmic maturation of the 60S ribosomal subunit. *Mol Cell* **39**: 196–208.

Longtine MS, McKenzie A 3rd, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast* **14**: 953–961.

López-Bigas N, Audit B, Ouzounis C, Parra G, Guigó R. 2005. Are splicing mutations the most frequent cause of hereditary disease? *FEBS Lett* **579**: 1900–1903.

Maeder C, Kutach AK, Guthrie C. 2009. ATP-dependent unwinding of U4/U6 snRNAs by the Brr2 helicase requires the C terminus of Prp8. *Nat Struct Mol Biol* **16**: 42–48.

Makarov EM, Makarova OV, Urlaub H, Gentzel M, Will CL, Wilm M, Lührmann R. 2002. Small Nuclear Ribonucleoprotein Remodeling During Catalytic Activation of the Spliceosome. *Science* **298**: 2205–2208.

Mallam AL, Campo MD, Gilman B, Sidote DJ, Lambowitz AM. 2012. Structural basis for RNA-duplex recognition and unwinding by the DEAD-box helicase Mss116p. *Nature* **490**: 121–125.

Martin A, Schneider S, Schwer B. 2002. Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. *J Biol Chem* **277**: 17743–17750.

Martin-Tumasz S, Reiter NJ, Brow DA, Butcher SE. 2010. Structure and functional implications of a complex containing a segment of U6 RNA bound by a domain of Prp24. *RNA* **16**: 792–804.

Martin-Tumasz S, Richie AC, Clos LJ, Brow DA, Butcher SE. 2011. A novel occluded RNA recognition motif in Prp24 unwinds the U6 RNA internal stem loop. *Nucleic Acids Res* **39**: 7837–7847.

Mattick JS, Makunin IV. 2006. Non-coding RNA. *Hum Mol Genet* **15(suppl 1)**: R17–29.

Mayas RM, Maita H, Staley JP. 2006. Exon ligation is proofread by the DExD/H-box ATPase Prp22p. *Nat Struct Mol Biol* **13**: 482–490.

McKie AB, McHale JC, Keen TJ, Tarttelin EE, Goliath R, van Lith-Verhoeven JJ, Greenberg J, Ramesar RS, Hoyng CB, Cremers FP, et al. 2001. Mutations in the pre-mRNA splicing factor gene PRPC8 in autosomal dominant retinitis pigmentosa (RP13). *Hum Mol Genet* **10**: 1555–1562.

Montpetit B, Thomsen ND, Helmke KJ, Seeliger MA, Berger JM, Weis K. 2011. A conserved mechanism of DEAD-box ATPase activation by nucleoporins and IP6 in mRNA export. *Nature* **472**: 238–242.

Mor A, Shav-Tal Y. 2010. Dynamics and kinetics of nucleo-cytoplasmic mRNA export. *Wiley Interdisciplinary Reviews: RNA* **1**: 388–401.

Morrissey JP, Tollervey D. 1993. Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. *Mol Cell Biol* **13**: 2469–2477.

Mouaikel J, Verheggen C, Bertrand E, Tazi J, Bordonné R. 2002. Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol Cell* **9**: 891–901.

Mougey EB, O'Reilly M, Osheim Y, Miller OL Jr, Beyer A, Sollner-Webb B. 1993. The terminal balls characteristic of eukaryotic rRNA transcription units in chromatin spreads are rRNA processing complexes. *Genes Dev* **7**: 1609–1619.

Nagasaki H, Arita M, Nishizawa T, Suwa M, Gotoh O. 2005. Species-specific variation of alternative splicing and transcriptional initiation in six eukaryotes. *Gene* **364**: 53–62.

Narla A, Ebert BL. 2010. Ribosomopathies: human disorders of ribosome dysfunction. *Blood* **115**: 3196–3205.

Niu Z, Jin W, Zhang L, Li X. 2012. Tumor suppressor RBM5 directly interacts with the DExD/H-box protein DHX15 and stimulates its helicase activity. *FEBS Letters* **586**: 977–983.

Noble KN, Tran EJ, Alcázar-Román AR, Hodge CA, Cole CN, Wente SR. 2011. The Dbp5 cycle at the nuclear pore complex during mRNA export II: nucleotide cycling and mRNP remodeling by Dbp5 are controlled by Nup159 and Gle1. *Genes Dev* **25**: 1065–1077.

Van Nues RW, Beggs JD. 2001. Functional contacts with a range of splicing proteins suggest a central role for Brr2p in the dynamic control of the order of events in spliceosomes of Saccharomyces cerevisiae. *Genetics* **157**: 1451–1467.

Oberer M, Marintchev A, Wagner G. 2005. Structural basis for the enhancement of eIF4A helicase activity by eIF4G. *Genes Dev* **19**: 2212–2223.

Oh JJ, Razfar A, Delgado I, Reed RA, Malkina A, Boctor B, Slamon DJ. 2006. 3p21.3 tumor suppressor gene H37/Luca15/RBM5 inhibits growth of human lung cancer cells through cell cycle arrest and apoptosis. *Cancer Res* **66**: 3419–3427.

Ooi SL, Samarsky DA, Fournier MJ, Boeke JD. 1998. Intronic snoRNA biosynthesis in Saccharomyces cerevisiae depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA. *RNA* **4**: 1096–1110.

Padgett RA, Konarska MM, Grabowski PJ, Hardy SF, Sharp PA. 1984. Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. *Science* **225**: 898–903.

Pagani F, Raponi M, Baralle FE. 2005. Synonymous mutations in CFTR exon 12 affect splicing and are not neutral in evolution. *Proc Natl Acad Sci USA* **102**: 6368–6372.

Pandit S, Lynn B, Rymond BC. 2006. Inhibition of a spliceosome turnover pathway suppresses splicing defects. *Proc Natl Acad Sci USA* **103**: 13700–13705.

Pandit S, Paul S, Zhang L, Chen M, Durbin N, Harrison SMW, Rymond BC. 2009. Spp382p Interacts with Multiple Yeast Splicing Factors, Including Possible Regulators of Prp43 DExD/H-Box Protein Function. *Genetics* **183**: 195–206.

Panse VG, Johnson AW. 2010. Maturation of eukaryotic ribosomes: acquisition of functionality. *Trends Biochem Sci* **35**: 260–266.

Parenteau J, Durand M, Morin G, Gagnon J, Lucier J-F, Wellinger RJ, Chabot B, Abou Elela S. 2011. Introns within Ribosomal Protein Genes Regulate the Production and Function of Yeast Ribosomes. *Cell* **147**: 320–331.

Parenteau J, Durand M, Véronneau S, Lacombe A-A, Morin G, Guérin V, Cecez B, Gervais-Bird J, Koh C-S, Brunelle D, et al. 2008. Deletion of Many Yeast Introns Reveals a Minority of Genes that Require Splicing for Function. *Mol Biol Cell* **19**: 1932–1941.

Patel SB, Bellini M. 2008. The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res* **36**: 6482–6493.

Peng J, Valeshabad AK, Li Q, Wang Y. 2013. Differential expression of RBM5 and KRAS in pancreatic ductal adenocarcinoma and their association with clinicopathological features. *Oncol Lett* **5**: 1000–1004.

Pérez-Fernández J, Román Á, De Las Rivas J, Bustelo XR, Dosil M. 2007. The 90S Preribosome Is a Multimodular Structure That Is Assembled through a Hierarchical Mechanism. *Mol Cell Biol* **27**: 5414–5429.

Perriman R, Barta I, Voeltz GK, Abelson J, Ares M Jr. 2003. ATP requirement for Prp5p function is determined by Cus2p and the structure of U2 small nuclear RNA. *Proc Natl Acad Sci USA* **100**: 13857–13862.

Pertschy B, Schneider C, Gnädig M, Schäfer T, Tollervey D, Hurt E. 2009. RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J Biol Chem* **284**: 35079–35091.

Phipps KR, Charette JM, Baserga SJ. 2011. The SSU Processome in Ribosome Biogenesis – Progress and Prospects. *Wiley Interdiscip Rev RNA* **2**: 1–21.

Pleiss JA, Whitworth GB, Bergkessel M, Guthrie C. 2007. Rapid, Transcript-Specific Changes in Splicing in Response to Environmental Stress. *Mol Cell* **27**: 928–937.

Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E, Bragado-Nilsson E, Wilm M, Séraphin B. 2001. The Tandem Affinity Purification (TAP) Method: A General Procedure of Protein Complex Purification. *Methods* **24**: 218–229.

Raghunathan PL, Guthrie C. 1998a. A Spliceosomal Recycling Factor That Reanneals U4 and U6 Small Nuclear Ribonucleoprotein Particles. *Science* **279**: 857 –860.

Raghunathan PL, Guthrie C. 1998b. RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. *Curr Biol* **8**: 847–855.

Raker VA, Hartmuth K, Kastner B, Lührmann R. 1999. Spliceosomal U snRNP Core Assembly: Sm Proteins Assemble onto an Sm Site RNA Nonanucleotide in a Specific and Thermodynamically Stable Manner. *Mol Cell Biol* **19**: 6554–6565.

Rearick D, Prakash A, McSweeny A, Shepard SS, Fedorova L, Fedorov A. 2011. Critical association of ncRNA with introns. *Nucleic Acids Res* **39**: 2357–2366.

Richards JD, Johnson KA, Liu H, McRobbie A-M, McMahon S, Oke M, Carter L, Naismith JH, White MF. 2008. Structure of the DNA repair helicase hel308 reveals DNA binding and autoinhibitory domains. *J Biol Chem* **283**: 5118–5126.

Rodriguez JR, Pikielny CW, Rosbash M. 1984. In vivo characterization of yeast mRNA processing intermediates. *Cell* **39**: 603–610.

Ruby SW, Chang TH, Abelson J. 1993. Four yeast spliceosomal proteins (PRP5, PRP9, PRP11, and PRP21) interact to promote U2 snRNP binding to pre-mRNA. *Genes Dev* **7**: 1909–1925.

Ruggero D, Pandolfi PP. 2003. Does the ribosome translate cancer? *Nature Reviews Cancer* **3**: 179–192.

Sakharkar MK, Chow VTK, Kangueane P. 2004. Distributions of exons and introns in the human genome. *In Silico Biol (Gedrukt)* **4**: 387–393.

Sambrook J, Russell DW. 2001. *Molecular cloning: a laboratory manual*. 3rd ed. CSHL Press.

Sanjuán R, Marín I. 2001. Tracing the Origin of the Compensasome: Evolutionary History of DEAH Helicase and MYST Acetyltransferase Gene Families. *Mol Biol Evol* **18**: 330–343.

Schäfer T, Maco B, Petfalski E, Tollervey D, Böttcher B, Aebi U, Hurt E. 2006. Hrr25-dependent phosphorylation state regulates organization of the pre-40S subunit. *Nature* **441**: 651–655.

Schneider BL, Steiner B, Seufert W, Futcher AB. 1996. pMPY-ZAP: a reusable polymerase chain reaction-directed gene disruption cassette for Saccharomyces cerevisiae. *Yeast* **12**: 129–134.

Schütz P, Bumann M, Oberholzer AE, Bieniossek C, Trachsel H, Altmann M, Baumann U. 2008. Crystal structure of the yeast eIF4A-eIF4G complex: an RNA-helicase controlled by protein-protein interactions. *Proc Natl Acad Sci USA* **105**: 9564–9569.

Schütz P, Karlberg T, van den Berg S, Collins R, Lehtiö L, Högbom M, Holmberg-Schiavone L, Tempel W, Park H-W, Hammarström M, et al. 2010. Comparative Structural Analysis of Human DEAD-Box RNA Helicases. *PLoS ONE* **5**: e12791.

Schwer B. 2008. A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. *Mol Cell* **30**: 743–754.

Schwer B, Gross CH. 1998. Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. *EMBO J* **17**: 2086–2094.

Schwer B, Guthrie C. 1992. A conformational rearrangement in the spliceosome is dependent on PRP16 and ATP hydrolysis. *EMBO J* **11**: 5033–5039.

Schwer B, Guthrie C. 1991. PRP16 is an RNA-dependent ATPase that interacts transiently with the spliceosome. *Nature* **349**: 494–499.

Seiser RM, Sundberg AE, Wollam BJ, Zobel-Thropp P, Baldwin K, Spector MD, Lycan DE. 2006. Ltv1 is required for efficient nuclear export of the ribosomal small subunit in Saccharomyces cerevisiae. *Genetics* **174**: 679–691.

Séraphin B, Kretzner L, Rosbash M. 1988. A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. *EMBO J* **7**: 2533–2538.

Seraphin B, Rosbash M. 1989. Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. *Cell* **59**: 349–358.

Shannon KW, Guthrie C. 1991. Suppressors of a U4 snRNA mutation define a novel U6 snRNP protein with RNA-binding motifs. *Genes Dev* **5**: 773–785.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**: 2498–2504.

Shimba S, Reddy R. 1994. Purification of human U6 small nuclear RNA capping enzyme. Evidence for a common capping enzyme for gamma-monomethyl-capped small RNAs. *J Biol Chem* **269**: 12419–12423.

Shiratori A, Shibata T, Arisawa M, Hanaoka F, Murakami Y, Eki T. 1999. Systematic identification, classification, and characterization of the open reading frames which encode novel helicase-related proteins in Saccharomyces cerevisiae by gene disruption and Northern analysis. *Yeast* **15**: 219–253.

Shulman RW, Warner JR. 1978. Ribosomal RNA transcription in a mutant of Saccharomyces cerevisiae defective in ribosomal protein synthesis. *Mol Gen Genet* **161**: 221–223.

Silverman EJ, Maeda A, Wei J, Smith P, Beggs JD, Lin R-J. 2004. Interaction between a G-Patch Protein and a Spliceosomal DEXD/H-Box ATPase That Is Critical for Splicing. *Mol Cell Biol* **24**: 10101–10110.

Sleeman JE, Lamond AI. 1999. Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. *Curr Biol* **9**: 1065–1074.

Small EC, Leggett SR, Winans AA, Staley JP. 2006. The EF-G-like GTPase Snu114p regulates spliceosome dynamics mediated by Brr2p, a DExD/H box ATPase. *Mol Cell* **23**: 389–399.

Spiller MP, Boon K-L, Reijns MAM, Beggs JD. 2007. The Lsm2-8 complex determines nuclear localization of the spliceosomal U6 snRNA. *Nucleic Acids Res* **35**: 923–929.

Spingola M, Grate L, Haussler D, Ares M. 1999. Genome-wide bioinformatic and molecular analysis of introns in Saccharomyces cerevisiae. *RNA* **5**: 221–234.

Staley JP, Guthrie C. 1999. An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p. *Mol Cell* **3**: 55–64.

Staley JP, Woolford JL. 2009. Assembly of ribosomes and spliceosomes: complex ribonucleoprotein machines. *Curr Opin Cell Biol* **21**: 109–118.

Stevens SW. 2010. Chapter 4 The Biology of DEAH/RHA Proteins and Their Mechanism of Action. In *RNA Helicases*, *RSC Biomolecular Series 19*, pp. 99–115, Royal Society of Chemistry, London.

Stevens SW, Abelson J. 1999. Purification of the yeast U4/U6.U5 small nuclear ribonucleoprotein particle and identification of its proteins. *Proc Natl Acad Sci USA* **96**: 7226–7231.

Stevens SW, Abelson J. 2002. Yeast pre-mRNA splicing: methods, mechanisms, and machinery. *Meth Enzymol* **351**: 200–220.

Stevens SW, Barta I, Ge HY, Moore RE, Young MK, Lee TD, Abelson J. 2001. Biochemical and genetic analyses of the U5, U6, and U4/U6 x U5 small nuclear ribonucleoproteins from Saccharomyces cerevisiae. *RNA* **7**: 1543–1553.

Stevens SW, Ryan DE, Ge HY, Moore RE, Young MK, Lee TD, Abelson J. 2002. Composition and functional characterization of the yeast spliceosomal penta-snRNP. *Mol Cell* **9**: 31–44.

Strunk BS, Loucks CR, Su M, Vashisth H, Cheng S, Schilling J, Brooks CL, Karbstein K, Skiniotis G. 2011. Ribosome Assembly Factors Prevent Premature Translation Initiation by 40S Assembly Intermediates. *Science* **333**: 1449–1453.

Strunk BS, Novak MN, Young CL, Karbstein K. 2012. A translation-like cycle is a quality control checkpoint for maturing 40S ribosome subunits. *Cell* **150**: 111–121.

Sung M-K, Ha CW, Huh W-K. 2008. A vector system for efficient and economical switching of C-terminal epitope tags in Saccharomyces cerevisiae. *Yeast* **25**: 301–311.

Tanaka N, Aronova A, Schwer B. 2007. Ntr1 activates the Prp43 helicase to trigger release of lariat-intron from the spliceosome. *Genes Dev* **21**: 2312–2325.

Tanaka N, Schwer B. 2005. Characterization of the NTPase, RNA-binding, and RNA helicase activities of the DEAH-box splicing factor Prp22. *Biochemistry* **44**: 9795–9803.

Tanaka N, Schwer B. 2006. Mutations in PRP43 that uncouple RNA-dependent NTPase activity and pre-mRNA splicing function. *Biochemistry* **45**: 6510–6521.

Tanner NK, Cordin O, Banroques J, Doère M, Linder P. 2003. The Q Motif: A Newly Identified Motif in DEAD Box Helicases May Regulate ATP Binding and Hydrolysis. *Molecular Cell* **11**: 127–138.

Tannukit S, Crabb TL, Hertel KJ, Wen X, Jans DA, Paine ML. 2009. Identification of a novel nuclear localization signal and speckle-targeting sequence of tuftelin-interacting protein 11, a splicing factor involved in spliceosome disassembly. *Biochem Biophys Res Commun* **390**: 1044–1050.

Tarn WY, Lee KR, Cheng SC. 1993. The yeast PRP19 protein is not tightly associated with small nuclear RNAs, but appears to associate with the spliceosome after binding of U2 to the pre-mRNA and prior to formation of the functional spliceosome. *Mol Cell Biol* **13**: 1883–1891.

The UniProt Consortium. 2011. Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Research* **40**: D71–D75.

Tsai R-T, Fu R-H, Yeh F-L, Tseng C-K, Lin Y-C, Huang Y, Cheng S-C. 2005. Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2. *Genes Dev* **19**: 2991–3003.

Tsai R-T, Tseng C-K, Lee P-J, Chen H-C, Fu R-H, Chang K, Yeh F-L, Cheng S-C. 2007. Dynamic Interactions of Ntr1-Ntr2 with Prp43 and with U5 Govern the Recruitment of Prp43 To Mediate Spliceosome Disassembly. *Mol Cell Biol* **27**: 8027–8037.

Tseng C-K, Cheng S-C. 2008. Both catalytic steps of nuclear pre-mRNA splicing are reversible. *Science* **320**: 1782–1784.

Tycowski KT, Shu MD, Steitz JA. 1996. A mammalian gene with introns instead of exons generating stable RNA products. *Nature* **379**: 464–466.

Umate P, Tuteja N, Tuteja R. 2011. Genome-wide comprehensive analysis of human helicases. *Commun Integr Biol* **4**: 118–137.

Ungar L, Yosef N, Sela Y, Sharan R, Ruppin E, Kupiec M. 2009. A genome-wide screen for essential yeast genes that affect telomere length maintenance. *Nucleic Acids Res* **37**: 3840–3849.

Uryu K, Nakashima-Yasuda H, Forman MS, Kwong LK, Clark CM, Grossman M, Miller BL, Kretzschmar HA, Lee VM-Y, Trojanowski JQ, et al. 2008. Concomitant TAR-DNA-binding protein 43 pathology is present in Alzheimer disease and corticobasal degeneration but not in other tauopathies. *J Neuropathol Exp Neurol* **67**: 555–564.

Venables JP, Koh C-S, Froehlich U, Lapointe E, Couture S, Inkel L, Bramard A, Paquet ER, Watier V, Durand M, et al. 2008. Multiple and specific mRNA processing targets for the major human hnRNP proteins. *Mol Cell Biol* **28**: 6033–6043.

Venema J, Henry Y, Tollervey D. 1995. Two distinct recognition signals define the site of endonucleolytic cleavage at the 5'-end of yeast 18S rRNA. *EMBO J* **14**: 4883–4892.

Venema J, Tollervey D. 1999. Ribosome synthesis in Saccharomyces cerevisiae. *Annu Rev Genet* **33**: 261–311.

Verheggen C, Bertrand E. 2012. CRM1 plays a nuclear role in transporting snoRNPs to nucleoli in higher eukaryotes. *Nucleus* **3**: 132–137.

Villa T, Guthrie C. 2005. The Isy1p component of the NineTeen complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing. *Genes Dev* **19**: 1894–1904.

Vithana EN, Abu-Safieh L, Allen MJ, Carey A, Papaioannou M, Chakarova C, Al-Maghtheh M, Ebenezer ND, Willis C, Moore AT, et al. 2001. A human homolog of yeast pre-mRNA splicing gene, PRP31, underlies autosomal dominant retinitis pigmentosa on chromosome 19q13.4 (RP11). *Mol Cell* **8**: 375–381.

Wahl MC, Will CL, Lührmann R. 2009. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* **136**: 701–718.

Walbott H, Mouffok S, Capeyrou R, Lebaron S, Humbert O, van Tilbeurgh H, Henry Y, Leulliot N. 2010. Prp43p contains a processive helicase structural architecture with a specific regulatory domain. *EMBO J* **29**: 2194–2204.

Warner JR. 1999. The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci* **24**: 437–440.

Warner JR, Udem SA. 1972. Temperature sensitive mutations affecting ribosome synthesis in Saccharomyces cerevisiae. *J Mol Biol* **65**: 243–257.

Watkins NJ, Bohnsack MT. 2012. The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdisciplinary Reviews: RNA* **3**: 397–414.

Weber G, Cristão VF, de L. Alves F, Santos KF, Holton N, Rappsilber J, Beggs JD, Wahl MC. 2011. Mechanism for Aar2p function as a U5 snRNP assembly factor. *Genes Dev* **25**: 1601–1612.

Wei M-H, Latif F, Bader S, Kashuba V, Chen J-Y, Duh F-M, Sekido Y, Lee C-C, Geil L, Kuzmin I, et al. 1996. Construction of a 600-Kilobase Cosmid Clone Contig and Generation of a Transcriptional Map Surrounding the Lung Cancer Tumor

Suppressor Gene (TSG) Locus on Human Chromosome 3p21.3: Progress Toward the Isolation of a Lung Cancer TSG. *Cancer Res* **56**: 1487–1492.

Weidenhammer EM, Ruiz-Noriega M, Woolford JL Jr. 1997. Prp31p promotes the association of the U4/U6 x U5 tri-snRNP with prespliceosomes to form spliceosomes in Saccharomyces cerevisiae. *Mol Cell Biol* **17**: 3580–3588.

Weir JR, Bonneau F, Hentschel J, Conti E. 2010. Structural analysis reveals the characteristic features of Mtr4, a DExH helicase involved in nuclear RNA processing and surveillance. *Proc Natl Acad Sci U S A* **107**: 12139–12144.

Wen X, Lei Y-P, Zhou YL, Okamoto CT, Snead ML, Paine ML. 2005. Structural organization and cellular localization of tuftelin-interacting protein 11 (TFIP11). *Cell Mol Life Sci* **62**: 1038–1046.

Wen X, Tannukit S, Paine ML. 2008. TFIP11 interacts with mDEAH9, an RNA helicase involved in spliceosome disassembly. *Int J Mol Sci* **9**: 2105–2113.

White J, Li Z, Sardana R, Bujnicki JM, Marcotte EM, Johnson AW. 2008. Bud23 methylates G1575 of 18S rRNA and is required for efficient nuclear export of pre-40S subunits. *Mol Cell Biol* **28**: 3151–3161.

Wiest DK, O'Day CL, Abelson J. 1996. In Vitro Studies of the Prp9·Prp11·Prp21 Complex Indicate a Pathway for U2 Small Nuclear Ribonucleoprotein Activation. *J Biol Chem* **271**: 33268–33276.

Wise JA. 1991. Preparation and analysis of low molecular weight RNAs and small ribonucleoproteins. *Meth Enzymol* **194**: 405–415.

Xu Y-Z, Query CC. 2007. Competition between the ATPase Prp5 and branch region-U2 snRNA pairing modulates the fidelity of spliceosome assembly. *Mol Cell* **28**: 838–849.

Yaffe MP, Schatz G. 1984. Two nuclear mutations that block mitochondrial protein import in yeast. *Proc Natl Acad Sci USA* **81**: 4819–4823.

Yang F, Wang X-Y, Zhang Z-M, Pu J, Fan Y-J, Zhou J, Query CC, Xu Y-Z. 2013. Splicing proofreading at 5′ splice sites by ATPase Prp28p. *Nucl Acids Res*. [Epub ahead of print].

Ye Y, De Leon J, Yokoyama N, Naidu Y, Camerini D. 2005. DBR1 siRNA inhibition of HIV-1 replication. *Retrovirology* **2**: 63.

Yeh T-C, Liu H-L, Chung C-S, Wu N-Y, Liu Y-C, Cheng S-C. 2011. Splicing Factor Cwc22 Is Required for the Function of Prp2 and for the Spliceosome To Escape from a Futile Pathway. *Mol Cell Biol* **31**: 43–53.

Yoshimoto R, Kataoka N, Okawa K, Ohno M. 2009. Isolation and characterization of post-splicing lariat-intron complexes. *Nucleic Acids Res* **37**: 891–902.

Zanchin NI, Goldfarb DS. 1999. The exosome subunit Rrp43p is required for the efficient maturation of 5.8S, 18S and 25S rRNA. *Nucleic Acids Res* **27**: 1283–1288.

Zhang Y, Lu H. 2009. Signaling to p53: ribosomal proteins find their way. *Cancer Cell* **16**: 369–377.

Zhou XZ, Lu KP. 2001. The Pin2/TRF1-Interacting Protein PinX1 Is a Potent Telomerase Inhibitor. *Cell* **107**: 347–359.

Vita

Jennifer Ann Hennigan was born in 1984 in Shreveport, Louisiana. She earned a B.S. in Biology from Louisiana Tech Univeristy in 2006 and graduated *summa cum laude*. Jennifer began graduate school at the University of Texas at Austin in the Department of Molecular Genetics and Microboiology in 2006. She currently resides with her husband of six years in Austin.

Permanent email: jennahennigan@gmail.com

This dissertation was typed by Jennifer Ann Hennigan.